利用數值資料產生模糊規則以解模糊分類問題之新方法
Generating Fuzzy Rules from Numerical Data for Handling Fuzzy Classification Problems

陳錫明*　李少華**　李嘉晃**
Shyi-Ming Chen, Shao-Hua Lee, and Chia-Hoang Lee

*國立台灣科技大學電子工程系
Department of Electronic Engineering
National Taiwan University of Science and Technology, Taipei, Taiwan, R. O. C.

**國立交通大學資訊科學系
Department of Computer and Information Science
National Chiao Tung University, Hsinchu, Taiwan, R. O. C.

摘要

模糊分類為模糊邏輯的一個重要應用。建造一個模糊分類系統的最重要工作為針對一個特定的分類問題去找出一組模糊規則以解此問題。本論文提出一個從數值資料產生模糊規則以解模糊分類問題的新方法。在本論文中，我們根據決策與屬性之各個語詞間的模糊子集合關係程度值，來決定決策與各個語詞間之密切度，從而藉由準位臨界值之設定，利用篩選而得出相關性高之語詞，以組合出該決策適用之模糊規則。本論文所提的方法比目前已存在之方法可產生較少的規則數，並有更高的分類準確率。

關鍵詞：模糊分類, 模糊規則, 模糊子集合關係程度值, 準位臨界值, 應用臨界值。

Abstract

Fuzzy classification is one of the important applications of fuzzy logic. The most important task to accomplish a fuzzy classification system is to find a set of fuzzy rules suitable for the specific classification problem. In this paper, we present a new method for generating fuzzy rules form numerical data for handling fuzzy classification problems based on the fuzzy subsethood values between decisions to be made and terms of attributes by using the level threshold value α and the applicability threshold value β, where $\alpha \in [0,1]$ and $\beta \in [0,1]$. The proposed method has a higher classification accuracy rate and generates fewer fuzzy rules than the existing methods.

Keywords: Fuzzy Classifications, Fuzzy Rules, Fuzzy Subsethood Values, Level Threshold Values, Applicabi lit Threshold Values.

1. Introduction

Fuzzy classification is one of the important applications of fuzzy logic [21], [22]. Fuzzy classification systems are capable of handling perceptual uncertainties, such as the vagueness and ambiguity involved in the classification problem [19]. The most important task to accomplish a fuzzy classification system is to find a set of fuzzy rules suitable for the specific classification problem. Usually, we have two methods to complete this task. One approach is to obtain knowledge from experts and translate their knowledge directly into fuzzy rules. However, the process of knowledge acquisition and validation is difficult and time-consuming. It is very likely that an expert may not be able to express his or her knowledge explicitly and accurately. Another approach is to generate fuzzy rules through a machine learning process [1], [3], [5], [7], [9], [14], [17], [19], [20], with which knowledge can be automatically extracted or induced from sample cases or examples. In [3], we have presented a method for generating fuzzy rules from relational database systems for estimating null values. In [17], we have presented a method for constructing membership functions and fuzzy rules from training examples.

A commonly used machine learning method is the induction of decision trees [16] for a specific problem. The method of decision trees induction has been expanded to induce fuzzy decision trees proposed by Yuan and Sha [19], where fuzzy entropy is used to lead the search of the most effective decision nodes. However, the method presented in [19] has some drawbacks, i.e., (1) It generates too many fuzzy rules. (2) Its classification accuracy rate is not good enough.

In this paper, we present a new method based on the filtering of the fuzzy subsethood values [12], [19] between decisions to be made and terms of attributes by the level threshold value α and the applicability threshold value β for generating fuzzy rules from the numerical data in a more efficient manner, where $\alpha \in [0,1]$ and $\beta \in [0,1]$. The proposed method has higher classification accuracy and generates fewer fuzzy rules than the one presented in [19].

This paper is organized as follows. In Section 2, the basi concepts of fuzzy sets are reviewed from [12], [19], and [21]. In Section 3, we briefly review Yuan-and-Shaw's fuzzy rules generation method from [19]. In Section 4, we propose a fuzzy learning method based on the fuzzy subsethood values between decisions to be made and terms of attributes b using the level threshold value α and the applicabilit threshold value β to directly generate fuzzy rules fro numerical data to deal with the fuzzy classification problem, where $\alpha \in [0,1]$ and $\beta \in [0,1]$. Furthermore, we also use the example shown in [19] to illustrate the fuzzy rules generation process. The conclusions are discussed in Section 5.

2. Fuzzy Set Theory

The theory of fuzzy sets was proposed by Zadeh in 1965 [21]. Roughly speaking, a fuzzy set is a set with fuzzy boundaries. A fuzzy set can be characterized by a membership function in a universe of discourse. A fuzzy set A in the universe of discourse U can be characterized b a membership function μ_A as follows

$$\mu_A : U \rightarrow [0, 1],$$

where the degree of membership $\mu_A(u)$ of an element u in the fuzzy set A is between zero and one and $u \in U$.

Definition 2.1: Let A and B be two fuzzy sets of the universe of discourse U with membership functions μ_A and μ_B, respectively. The union of the fuzzy sets A and B is defined b

$$\mu_{A \cup B}(u) = \max\{ \mu_A(u), \ \mu_B(u)\}, \quad \forall u \in U. \quad (1)$$

The intersection of A and B, $\cap$ B, is defined b

$$\mu_{A \cap B}(u) = \min\{ \mu_A(u), \ \mu_B(u)\}, \quad \forall u \in U. \quad (2)$$

The complement of A, denoted as $\overline{A}$, is defined b

$$\mu_{\overline{A}}(u) = 1 - \mu_A(u), \quad \forall u \in U. \quad (3)$$

Definition 2.2: Let A and B be two fuzzy sets of the universe of discourse U. Then, A is a subset of B if and only if

$$\mu_A(u) \leq \mu_B(u) \text{ for all } u \in U. \quad (4)$$

Definition 2.3: Let A and B be two fuzzy sets defined on the universe of discourse U with membership functions μ_A and μ_B, respectively. The fuzzy subsethood S(A, B [12], [19] measures the degree in which A is a subset of B

$$S(A, B) = \frac{M(A \cap B)}{M(A)} = \frac{\sum_{u \in U} Min(\mu_A(u), \mu_B(u))}{\sum_{u \in U} \mu_A(u)}, \quad (5)$$

where $S(A, B) \in [0, 1]$.

3. A Review of Yuan-and-Shaw's Method [19] for Fuzzy Rules Generation

In a fuzzy classification problem, a collection of cases U = {u} is represented by a set of attributes $A = \{A_1, \ldots, A_k\}$, where U is called the object space [19]. Each attribute A_k depicts some important feature of a case and is usuall limited to a small set of discrete linguistic terms $T(\ \ _k) = \{T_1^k, \ldots, T_{s_k}^k\}$. In other words, $T(A_k)$ is the domain of the attribute A_k. Each case u in U is classified into a class C_i, where C_i is a member of classes C and $C = \{C_1, \ldots, C_L\}$. In our discussions, both cases and classes are fuzzy. The class C_i of C, i = 1, ..., L, is a fuzzy set defined on the universe of cases U. The membership function $\mu_{C_i}(u)$ assigns a degree to which u belongs to class C_i. The attribute $\ \ _k$ is a linguistic variable which takes linguistic values from $T(\ \ _k) = \{ T_1^k, \ldots, T_{s_k}^k \}$. The linguistic values $\ \ _j^k$ are als fuzzy sets defined on U. The membership value $\mu_{T_j^k}(u)$ depicts the degree to which case u's attribute A_k is T_j^k. A fuzzy classification rule (or abbreviated into fuzzy rule) can be written in the form

$$\textit{IF } (A_1 \text{ is } T_{i_1}^1) \textit{ AND} \ldots \textit{AND } (A_k \text{ is } T_{i_k}^k) \textit{ THEN } (C \text{ is } C_j). \quad (6)$$

Using a machine learning method from a training set of cases whose class is known can induce a set of classification rules. An example of a small training data set of the Saturday Morning problem [19] with fuzzy membership values is shown in Table 1. In the Saturda Morning Problem, a case is a Saturday morning' s weather which can have four attributes

Attribute = {Outlook, Temperature, Humidity, Wind},

and each attribute has linguistic values

Outlook = {Sunny, Cloudy, Rain},
Temperature = {Hot, Mild, Cool},
Humidity = {Humid, Normal},
Wind = {Windy, Not-windy}.

The classification result (i.e., Plan) is the sport to be taken on that weekend day,

Plan = {Volleyball, Swimming, Weight-lifting}.

The fuzzy decision tree induction method presented in [19] consists of the following steps:

(1) Fuzzification of the training data.

(2) Induction of a fuzzy decision tree.

(3) Conversion of the decision tree into a set of rules.

(4) Application of the fuzzy rules for classification.

Using the data shown in Table 1, the generated fuzzy decision tree is shown in Fig. 1. From the fuzzy decision tree shown in Fig. 1, we can enumerate the number of routes from root to leaf. Each route can be converted into a rule, where the condition part represents the attributes on the passing branches from the root to the leaf and the conclusion pa rt represents the class at the leaf with the conclusion part represents the class at the leaf with the highest classification truth level. The generated fuzzy rules after conversion from the fuzzy decision tree are also shown in Fig. 1. In [19], Yuan et al. pointed out that Rule 3': "*IF Temperature* is *Hot AND Outlook* is *Rain THEN Weight-lifting*" can be simplified into Rule 3': "*IF Outlook* is *Rain THEN Weight-lifting*". The truth level of Rule 3' is 0.89 and is not less than 0.73 (the truth level of the origi nal Rule 3). With the generated six fuzzy rules shown in Fig. 1, the classification results for the training data shown in Table 1 are calculated. Among sixteen training cases, thirteen cases (except cases 2, 8, 16) are correctly classified The classification accuracy of the Yuan-and Shaw's method is 81%. For more details, please refer to [19].

4. A New Method for Generating Fuzzy Rules from Numerical Data

Table 1
A Small Data Set for the Saturday Morning Problem [19]

	Outlook			Temperature			Humidit		Wind		Plan		
Case	Sunny	Cloudy	Rain	Hot	Mild	Cool	Humid	Normal	Windy	Not-windy	Volleyball	Swimming	W-liftin
1	0.9	0.1	0	1	0	0	0.8	0.2	0.4	0.6	0	0.8	0.2
2	0.8	0.2	0	0.6	0.4	0	0	1	0	1	1	0.7	0
3	0	0.7	0.3	0.8	0.2	0	0.1	0.9	0.2	0.8	0.3	0.6	0.1
4	0.2	0.7	0.1	0.3	0.7	0	0.2	0.8	0.3	0.7	0.9	0.1	0
5	0	0.1	0.9	0.7	0.3	0	0.5	0.5	0.5	0.5	0	0	1
6	0	0.7	0.3	0	0.3	0.7	0.7	0.3	0.4	0.6	0.2	0	0.8
7	0	0.3	0.7	0	0	1	0	1	0.1	0.9	0	0	1
8	0	1	0	0	0.2	0.8	0.2	0.8	0	1	0.7	0	0.3
9	1	0	0	1	0	0	0.6	0.4	0.7	0.3	0.2	0.8	0
10	0.9	0.1	0	0	0.3	0.7	0	1	0.9	0.1	0	0.3	0.7
11	0.7	0.3	0	1	0	0	1	0	0.2	0.8	0.4	0.7	0
12	0.2	0.6	0.2	0	1	0	0.3	0.7	0.3	0.7	0.7	0.2	0.1
13	0.9	0.1	0	0.2	0.8	0	0.1	0.9	1	0	0	0	1
14	0	0.9	0.1	0	0.9	0.1	0.1	0.9	0.7	0.3	0	0	1
15	0	0	1	0	0	1	1	0	0.8	0.2	0	0	1
16	1	0	0	0.5	0.5	0	0	1	0	1	0.8	0.6	0

A. <u>Fuzzy decision tree</u>

Temperature? (G (Temperature) = 0.48)

Hot (G (Hot) = 0.45): Outlook? (G (Outlook | Hot) = 0.42)

Sunny: Swimming (S = 0.85)

Cloudy: Swimming (S = 0.72)

Rain: Weight-lifting (S = 0.73)

Mild (G (Mild) = 0.83): Wind? (G (Wind | Mild) = 0.36)

Windy: Weight-lifting (S = 0.81)

Not-windy: Volleyball (S = 0.78)

Cool(G (Cool) = 0.20): Weight-lifting (S = 0.88)

Note: G is the classification ambiguity measure at the decision node.
S is the classification truth level at the leaf.

B. <u>Fuzzy rules converted from the fuzzy decision tree</u>

Rule 1: ***IF*** Temperature is Hot ***AND*** Outlook is Sunny ***THEN*** Swimming (S = 0.85)

Rule 2: ***IF*** Temperature is Hot ***AND*** Outlook is Cloudy ***THEN*** Swimming (S = 0.72)

Rule 3: ***IF*** Temperature is Hot ***AND*** Outlook is Rain ***THEN*** Weight-lifting (S = 0.73)

Rule 4: ***IF*** Temperature is Mild ***AND*** Wind is Windy ***THEN*** Weight-lifting (S = 0.81)

Rule 5: ***IF*** Temperature is Mild ***AND*** Wind is Not-windy ***THEN*** Volleyball (S = 0.81)

Rule 6: ***IF*** Temperature is Cool ***THEN*** Weight-lifting (S = 0.88)

Note: Rule 3 can be simplified to Rule 3':

Rule 3': ***IF*** Outlook is Rain ***THEN*** Weight-lifting (S = 0.89)

Fig. 1. The induced fuzzy decision tree and fuzzy rules of Yuan-and-Shaw's method [19].

In this section, we present a new method for generatin fuzzy rules from numerical data. The data set we use to introduce the concepts of fuzzy rules generation is sh own in Table 2. In Table 2, we have nine cases with three attributes for each case and three kinds of decisions for each plan

Attribute = {A, B, C},

and each attribute has linguistic terms

A = {A1, A2, A3},
B = {B1, B2, B3 },
C = {C1, C2}.

The classification is the decision to be made on a case with

Table 2
A Data Set for Illustrating the Proposed Fuzzy Rules Generation Method

Case	A			B			C		Plan		
	A1	A2	A3	B1	B2	B3	C1	C2	X	Y	Z
1	0.3	0.7	0	0.2	0.7	0.1	0.3	0.7	0.1	0.9	0
2	1	0	0	1	0	0	0.7	0.3	0.8	0.2	0
3	0	0.3	0.7	0	0.7	0.3	0.6	0.4	0	0.2	0.8
4	0.8	0.2	0	0	0.7	0.3	0.2	0.8	0.6	0.3	0.1
5	0.5	0.5	0	1	0	0	0	1	0.6	0.8	0
6	0	0.2	0.8	0	1	0	0	1	0	0.7	0.3
7	1	0	0	0.7	0.3	0	0.2	0.8	0.7	0.4	0
8	0.1	0.8	0.1	0	0.9	0.1	0.7	0.3	0	0	1
9	0.3	0.7	0	0.9	0.1	0	1	0	0	0	1

attributes Ai, Bj, and Ck, respectively, to carry out one of the plans X, Y or Z:

$$\text{Plan} = \{X, Y, Z\}.$$

We want to generate fuzzy classification rules from the given numerical data in Table 2. The generated fuzzy classification rules are in the form of formula (6).

As shown in Table 2, we have nine cases, where each case has three attributes to describe it. For each attribute, we have two or three terms to choose. In addition to the attribute part, we have to decide on a plan. One of decisions "X", "Y" or "Z" is the plan to be decided for a specific case The value accompanying each term or decision of plan is in the range [0, 1]. For each case, we can decide which decision of plan (with the highest possibility value) is mos t likely to be chosen. For example, in Case 6, the possibility to choose decision "X" is 0, to choose decision "Y" is 0.7, to choose decision "Z" is 0.3, and the final decision is plan "Y".

From the possibility values of decisions "X", "Y", and "Z", for each case, we can decide which decision to be made for a specific case. If we divide the nine cases int three subgroups according to the classification results, i.e., "X", "Y", and "Z", we can get another table as shown in Table 3. As Table 3 depicted, there are three instances for "X", three instances for "Y", and three instances for "Z", respectively. After carefully examining the table, it seems that there are close relationships between classification results (decision of plan for that subgroup) and some terms of the attributes. Making use of the fuzzy subsethood concept [12], [19], we can get information about the relationship between the decision of the plan and every distinct term of the attributes.

Table 3
Three Subgroups According to the Decision to be Made

Subgroup	Case	A			B			C		Plan		
		A1	A2	A3	B1	B2	B3	C1	C2	X	Y	Z
Subgroup_1	2	1	0	0	1	0	0	0.7	0.3	0.8	0.2	0
	4	0.8	0.2	0	0	0.7	0.3	0.2	0.8	0.6	0.3	0.1
	7	1	0	0	0.7	0.3	0	0.2	0.8	0.7	0.4	0
Subgroup_2	1	0.3	0.7	0	0.2	0.7	0.1	0.3	0.7	0.1	0.9	0
	5	0.5	0.5	0	1	0	0	0	1	0.6	0.8	0
	6	0	0.2	0.8	0	1	0	0	1	0	0.7	0.3
Subgroup_3	3	0	0.3	0.7	0	0.7	0.3	0.6	0.4	0	0.2	0.8
	8	0.1	0.8	0.1	0	0.9	0.1	0.7	0.3	0	0	1
	9	0.3	0.7	0	0.9	0.1	0	1	0	0	0	1

In each subgroup, we calculate the fuzzy subsethood values between decisions of that subgroup and every term of each attribute. After the computations of subsethood values, we can get a set of subsethood values for each decision. In this set of values, the larger the value, the closer the relationship between the decision of the plan and the term. For each subgroup, we can attain the most important factors that result in the decision of the plan of that subgroup. We can use these terms to form the condition part of the classification rule for that decision of the plan. The consequent part of the rule is the decision of the plan for that subgroup.

From Table 3, we can see that there are three subgroups of cases. In each subgroup, the decision to be made is fixed. To find the closeness between the deci sion and each term of the three attributes, we first calculate the subsethood values for them. The meaning of fuzzy subsethood value is defined by using formula (5), A is a subset of B, defined b

$$S(A, B) = \frac{M(A \cap B)}{M(A)}.$$

Take Subgroup_1 as an example ("X" is the decision of Subgroup_1), the denominator and the numerator of the subsethood formula for S(X, A1) are as follows

$$M(X) = 0.8 + 0.6 + 0.7 = 2.1,$$
$$\begin{aligned} M(X \cap A1) &= \text{Min}(0.8, 1) + \text{Min}(0.6, 0.8) + \text{Min}(0.7, 1) \\ &= 0.8 + 0.6 + 0.7 \\ &= 2.1. \end{aligned}$$

The value of S(X, A1) is

$$\begin{aligned} S(X, A1) &= M(X \cap A1)/M(X) \\ &= 2.1/2.1 \\ &= 1, \end{aligned}$$

where S(X, A1) stands for the subsethood of "X" to "A1" of "A" in Subgroup_1.

Using the same formula (i.e., formula (5)), we can compute all the subsethood values as summarized in Fig. 2. From Fig. 2, we can find that some terms are closel related to the decision to be made in that subgroup and some are not. We need a standard to distinguish close or not close enough between the decision and terms of attributes. We use the level threshold value α as the standard to measure close enough or not on fuzzy subsethood values between the decision of the subgroup and all terms of attributes, where $\alpha \in [0, 1]$. Assume that the value we assigned to the level threshold α is 0.9. For each attribute, we can select at most one term. If there are two or more terms belonging to the same attribute which have a fuzzy subsethood value not less than 0.9, the one with the largest fuzzy subsethood value will be chosen. If there are two terms with subsethood values not less than 0.9 at the same time, the term which is the original term of the attribute will have privilege over the one which is a complemented term of the same attribute.

Subgroup_1(X):
A:
S(X, A1) = 1   S(X, A2) = 0.1   S(X, A3) = 0
B:
S(X, B1) = 0.71   S(X, B2) = 0.43   S(X, B3) = 0.14
C:
S(X, C1) = 0.52   S(X, C2) = 0.76
Subgroup_2(Y):
A:
S(Y, A1) = 0.33   S(Y, A2) = 0.58   S(Y, A3) = 0.29
B:
S(Y, B1) = 0.42   S(Y, B2) = 0.58   S(Y, B3) = 0.04
C:
S(Y, C1) = 0.13   S(Y, C2) = 0.92
Subgroup_3(Z):
A:
S(Z, A1) = 0.14   S(Z, A2) = 0.64   S(Z, A3) = 0.29
B:
S(Z, B1) = 0.32   S(Z, B2) = 0.61   S(Z, B3) = 0.14
C:
S(Z, C1) = 0.82   S(Z, C2) = 0.25

Fig. 2. The list of the f uzzy subsethood values for small data set.

Referring to Fig. 2, the fuzzy subsethood values (includin those for the complemented terms) not less than the level threshold value α, where $\alpha = 0.9$, in Subgroup_1 are S(X, A1) = 1, S(X, *NOT* A2) = 0.9 and S(X, *NOT* A3) = 1. Because "A1", "A2" and "A3" are all terms of attribute "A", only one of them will be chosen. In this condition, "A1" is the only original term that belongs to attribute "A", and it is the one we choose among them. From this term we can generate the first fuzzy rule as follows

Rule 1: *IF* A is A1 *THEN* Plan is X.

Likewise, the fuzzy subsethood values that are not less than 0.9 in Subgroup_2 are S(Y, *NO* B3) = 0.96 and S(Y, C2) = 0.92. The generated fuzzy rule is as follows

Rule 2: *IF* B is *NOT* B3 *AND* C is C2 *THEN* Plan is Y.

From Subgroup_3, we can see that the subsethood values are quite average. In this condition, no term is outstandin among them (no term has a value not less than 0.9). This means that for decision "Z", those terms of attri butes are average and no terms are representative enough. Thus, Rule 3 is unable to be generated at this time.

We use MF(Rule i) = MF(condition part of Rule i), where $1 \leq i \leq 2$, and MF means membership function value [19]. If we want to classify Case 3 of Table 3, then we can get

MF(condition part of Rule 1) = MF(A1) = 0,
MF(condition part of Rule 2) = MF(*NOT* B3 $\cap$ C2)
= $(1 - 0.3) \cap 0.4 = 0.4$,
MF(Rule 1) = MF(condition part of Rule 1) = 0,
MF(Rule 2) = MF(condition part of Rule 2) = 0.4.

Because both membership values of the existing rules are not high enough to choose decision "X" or decision "Y", it is very possible that decision "Z" is more appropriate than the other two decisions. In this situation, we need another applicability threshold value β, where $\beta \in [0, 1]$, to judge the applicability of the existing rules. The existing rules are applicable to a case if MF(Rule i) $\geq \beta$, where $i \in \{1, \ldots, n$ and n is the number of existing rules.

As an alternative, we can conclude that a case that is not well classified by Rule 1 and Rule 2 will be classified into the plan with decision "Z". Thus, the third fuzzy rule is generated as follows

Rule 3: *IF* MF(Rule 1) $< \beta$ *AND* MF(Rule 2) $< \beta$ *THEN* Plan is Z,

where MF(Rule i) = MF(con dition part of Rule i), where $1 \leq i \leq 2$, and MF means membership value [19], and β is a applicability threshold value that MF(Rule 1) or MF(Rule 2) must exceed if that rule is applicable to a case, where $\beta \in [0, 1]$. For Case 3, assume that the applicabilit y threshold value β is 0.6, then we can get

MF(condition part of Rule 1) = MF(A1) = 0,
MF(condition part of Rule 2) = MF(*NOT* B3 $\cap$ C2)
= $(1 - 0.3) \cap 0.4 = 0.4$,
MF(Rule 1) = MF(condition part of Rule 1) = 0,
MF(Rule 2) = MF(condition part of Rule 2) = 0.4.

Because MF(Rule 1) $< \beta$ and MF(Rule 2) $< \beta$, where $\beta = 0.6$, and according to Rule 3, we can see that the decision t be made for Case 3 is plan "Z".

To apply the generated fuzzy rules to each case of the data set shown in Table 2, we must assign the applicabilit threshold value β in advance, where $\beta \in [0, 1]$. For each case, calculate MF(Condition part of Rule 1) and MF(Condition part of Rule 2), respectively, and then assign MF(Rule 1) = MF(Condition part of Rule 1), MF(Ru le 2) = MF(Condition part of Rule 2). The applicability threshold value β is used to compare MF(Rule 1) and MF(Rule 2), respectively, for the specified case. If both MF(Rule 1) and MF(Rule 2) are less than β, then we let MF(Rule 3) = 1. Otherwise, we let MF(Rule 3) = 0.

In the example of Table 2, the classification results of Rule 1, Rule 2, and Rule 3 are "X", "Y", and "Z", respectively. The possibility values of the classification result for a specific case with respect to "X", "Y", and "Z" are represented by "Plan(X)", "Plan(Y)", and "Plan(Z)", respectively. After the calculations of MF(Rule 1), MF(Rule 2), and MF(Rule 3) for a specific case, we can assign

Plan(X) = MF(Rule 1),
Plan(Y) = MF(Rule 2), (7)
Plan(Z) = MF(Rule 3).

The generated fuzzy rules at the level threshold value α = 0.9 are listed as follows

Rule 1: *IF* A is A1 *THEN* Plan is X.
Rule 2: *IF B* is *NOT* B3 *AND* C is C2 *THEN* Plan is Y.
Rule 3: *IF* MF(Rule 1) $< \beta$ *AND* MF(Rule 2) $< \beta$ *THEN* Plan is Z.

Assume that the applicability threshold value β in the explained example is 0.6 (i.e., $\beta = 0.6$), then

(1) From Case 1 of Table 2, we can get
MF(condition part of Rule 1) = MF(A is A1) = 0.3,
MF(condition part of Rule 2) = MF(B is *NOT* B3 *AND* C is C2)
= MF(B is *NO* B3 $\cap$ C is C2)
= MF(B is *NO* B3) $\cap$ MF(C is C2)
= Min{(1 – 0.1), 0.7}
= 0.7,
MF(Rule 1) = MF(condition part of Rule 1) = 0.3,
MF(Rule 2) = MF(condition part of Rule 2) = 0.7.
Because MF(Rule 1) $< \beta$ and MF(Rule 2) $> \beta$, where $\beta = 0.6$, thus MF(Rule 3) = 0.
From formula (7), the possibility values of the decisions of plan for Case 1 are
Plan(X) = MF(Rule 1) = 0.3,
Plan(Y) = MF(Rule 2) = 0.7,
Plan(Z) = MF(Rule 3) = 0,
and we fill Plan(X), Plan(Y), and Plan(Z) (i.e., 0.3, 0.7, 0) into the last three columns of Case 1 in Table 4. Because Plan(Y) is the one with the highest possibility value among the values of P lan(X), Plan(Y), and Plan(Z), the decision to be made for Case 1 is "Y".

(2) From Case 2 of Table 2, we can get
MF(condition part of Rule 1) = MF(A is A1) = 1,
MF(condition part of Rule 2) = MF(B is *NOT* B3 *AND* C is C2)
= MF(B is *NOT* B3 $\cap$ C is C2)
= MF(B is *NOT* B3) $\cap$ MF(C is C2)
= Min{(1 – 0), 0.3}
= 0.3,

MF(Rule 1) = MF(condition part of Rule 1) = 1,
MF(Rule 2) = MF(condition part of Rule 2) = 0.3.
Because MF(Rule 1) > β and MF(Rule 2) < β, where β = 0.6, thus MF(Rule 3) = 0.
From formula (7), the possibility values of the decisions of plan for Case 2 are
Plan(X) = MF(Rule 1) = 1,
Plan(Y) = MF(Rule 2) = 0.3,
Plan(Z) = MF(Rule 3) = 0,
and we fill Plan(X), Plan(Y), and Plan(Z) (i.e., 1, 0.3, 0) into the last three columns of Case 2 in Table 5.
Because Plan(X) is the one with the highest possibilit value among the values of Plan(X), Plan(Y), and Plan(Z), the decision to be made for Case 2 is "X".

The other cases are treated in a similar way. We summarize the result in Table 4.

Based on the generated fuzzy classification rules, the classification results for the training data in Table 2 are shown in Table 4. Among nine training cases, all cases are correctly classified. The classification accuracy rate is 100%.

Table 4

Results after Applying the Generated Fuzzy Rules to Table 3

Case	A			B			C		Plan		
	A1	A2	A3	B1	B2	B3	C1	C2	X	Y	Z
1	0.3	0.7	0	0.2	0.7	0.1	0.3	0.7	0.3	0.7	0
2	1	0	0	1	0	0	0.7	0.3	1	0.3	0
3	0	0.3	0.7	0	0.7	0.3	0.6	0.4	0	0.4	1
4	0.8	0.2	0	0	0.7	0.3	0.2	0.8	0.8	0.7	0
5	0.5	0.5	0	1	0	0	0	1	0.5	1	0
6	0	0.2	0.8	0	1	0	0	1	0	1	0
7	1	0	0	0.7	0.3	0	0.2	0.8	1	0.8	0
8	0.1	0.8	0.1	0	0.9	0.1	0.7	0.3	0.1	0.3	1
9	0.3	0.7	0	0.9	0.1	0	1	0	0.3	0	1

In the following, we use an example [19] (i.e., the Saturday Morning Problem) to illustrate the fuzzy rules generation process.

Example 4.1: Assume that the small data set we use here is the same as [19] and shown in Table 1. From Table 1, we can see that there are four attributes for each case and there are three kinds of sport for each plan

Attribute = {Outlook, Temperature, Humidity, Wind},
and each attribute has terms shown as follows
Outlook = {Sunny, Cloudy, Rain},
Temperature = {Hot, Cool, Mild},
Humidity = {Humid, Normal},
Wind = {Windy, Not-windy}.
The classification result is the sport plan to be played on the weekend day:
Plan = {Volleyball, Swimming, Weight-lifting}.

Assume that the values for level threshold value α and applicability threshold value β are 0.9 and 0.6, respectively,(i.e., α = 0.9 and β = 0.6). From Table 1, w divide the sixteen cases into three subgroups according to the sport plan with the highest possibility value in each case The result of the division is as follows (refer to Table 5):

(1) Subgroup_1 with "Volleyball" as the activity to be taken Cases 2, 4, 8, 12, and 16.
(2) Subgroup_2 with "Swimming" as the activity to be taken: Cases 1, 3, 9, and 11.
(3) Subgroup_3 with "Weight-lifting" as the activity to be taken: Cases 5, 6, 7, 10, 13, 14, and 15.

According to formula (5), the calculations for subsethood for all three subgroups are shown in Fig. 3 for the Saturday Morning Problem.

According to the previous discussions, there are three fuzzy rules to be generated for α = 0.9 and β = 0.6 which are summarized as follows

Rule 1: *IF* Outlook is *NOT* Rain *AND* Humidity is Normal *AND* Wind is Not-windy *THEN* Plan is Volleyball.
Rule 2: *IF* Outlook is *NOT* Rain *AND* Temperature is Hot *THEN* Plan is Swimming.
Rule 3: *IF* MF(Rule1) < β *AND* MF(Rule2) < β *THEN* Plan is Weight-lifting.

Based on the previous discussions, we can apply the generated fuzzy rules to Table 1, The classification results of the application of the generated fuzzy rules are summarized in Table 6. From Table 6, we can see that among sixteen training cases, fifteen cases (except Case 3) are correctly classified. The classification accuracy rate is

$$\frac{15}{16} \times 100\% = 93.75\%.$$

A comparison of the number of generated fuzzy rules and accuracy rate between the proposed method and Yuan-and-Shaw's method [19] is listed in Table 7. From Table 7,

Table 5

Three Subgroups According to the Sport to be Taken

Sub-group	Case	Outlook			Temperature			Humidity		Wind		Plan		
		Sunny	Cloudy	Rain	Hot	Mild	Cool	Humid	Normal	Windy	Not-windy	Volleyball	Swimming	W-lifting
Subgroup_1	2	0.8	0.2	0	0.6	0.4	0	0	1	0	1	1	0.7	0
	4	0.2	0.7	0.1	0.3	0.7	0	0.2	0.8	0.3	0.7	0.9	0.1	0
	8	0	1	0	0	0.2	0.8	0.2	0.8	0	1	0.7	0	0.3
	12	0.2	0.6	0.2	0	1	0	0.3	0.7	0.3	0.7	0.7	0.2	0.1
	16	1	0	0	0.5	0.5	0	0	1	0	1	0.8	0.6	0
Subgroup_2	1	0.9	0.1	0	1	0	0	0.8	0.2	0.4	0.6	0	0.8	0.2
	3	0	0.7	0.3	0.8	0.2	0	0.1	0.9	0.2	0.8	0.3	0.6	0.1
	9	1	0	0	1	0	0	0.6	0.4	0.7	0.3	0.2	0.8	0
	11	0.7	0.3	0	1	0	0	1	0	0.2	0.8	0.4	0.7	0
Subgroup_3	5	0	0.1	0.9	0.7	0.3	0	0.5	0.5	0.5	0.5	0	0	1
	6	0	0.7	0.3	0	0.3	0.7	0.7	0.3	0.4	0.6	0.2	0	0.8
	7	0	0.3	0.7	0	0	1	0	1	0.1	0.9	0	0	1
	10	0.9	0.1	0	0	0.3	0.7	0	1	0.9	0.1	0	0.3	0.7
	13	0.9	0.1	0	0.2	0.8	0	0.1	0.9	1	0	0	0	1

	14	0	0.9	0.1	0	0.9	0.1	0.1	0.9	0.7	0.3	0	0	1
	15	0	0	1	0	0	1	1	0	0.8	0.2	0	0	1

we can see that the accuracy rate of the proposed method is better than that of Yuan-and-Shaw's method under $\alpha = 0.9$ and $\beta = 0.6$. The number of rules generated by the proposed method is less than the number of rules generated by Yuan-and-Shaw's method.

Subgroup_1 (Volleyball) :
Outlook:
S(Volleyball, Sunny) = 0.49 S(Volleyball, Cloudy) = 0.54
S(Volleyball, Rain) = 0.07
Temperature:
S(Volleyball, Hot) = 0.34 S(Volleyball, Mild) = 0.61
S(Volleyball, Cool) = 0.17
Humidity:
S(Volleyball, Humid) = 0.17 S(Volleyball, Normal) = 0.98
Wind:
S(Volleyball, Windy) = 0.15 S(Volleyball, Not-windy) = 0.95
Subgroup_2 (Swimming) :
Outlook:
S(Swimming, Sunny) = 0.79 S(Swimming, Cloudy) = 0.35
S(Swimming, Rain) = 0.10
Temperature:
S(Swimming, Hot) = 1 S(Swimming, Mild) = 0.07
S(Swimming, Cool) = 0
Humidity:
S(Swimming, Humid) = 0.76 S(Swimming, Normal) = 0.41
Wind:
S(Swimming, Windy) = 0.52 S(Swimming, Not-windy) = 0.76
Subgroup_3 (Weight-lifting) :
Outlook:
S(Weight-lifting, Sunny) = 0.25 S(Weight-lifting, Cloudy) = 0.34
S(Weight-lifting, Rain) = 0.46
Temperature:
S(Weight-lifting, Hot) = 0.14 S(Weight-lifting, Mild) = 0.4
S(Weight-lifting, Cool) = 0.54
Humidity:
S(Weight-lifting, Humid) = 0.37 S(Weight-lifting, Normal) = 0.66
Wind:
S(Weight-lifting, Windy) = 0.65 S(Weight -lifting, Not-windy) = 0.4

Fig. 3. The list of the fuzzy subsethood values.

Table 6
Learning Result of the Saturday Morning Problem with Generated Fuzzy Rules

Case	Classification Known in Training Data			Classification with Learned Rules		
	Volleyball	Swimming	W-lifting	Volleyball	Swimming	W-lifting
1	0.0	0.8	0.2	0.2	1	0
2	1.0	0.7	0.0	1	0.6	0
3	0.3	0.6	0.1	0.7	0.7	0[b]
4	0.9	0.1	0.0	0.7	0.3	0
5	0.0	0.0	1.0	0.1	0.1	1
6	0.2	0.0	0.8	0.3	0	1
7	0.0	0.0	1.0	0.3	0	1
8	0.7	0.0	0.3	0.8	0	0
9	0.2	0.8	0.0	0.3	1	0
10	0.0	0.3	0.7	0.1	0	1
11	0.4	0.7	0.0	0	1	0
12	0.7	0.2	0.1	0.7	0	0
13	0.0	0.0	1.0	0	0.2	1
14	0.0	0.0	1.0	0.3	0	1
15	0.0	0.0	1.0	0	0	1
16	0.8	0.6	0.0	1	0.5	0

a Wrong classification.
b Cannot distinguish between two or more classes.

Table 7
A Comparison of the Number of Generated Fuzzy Rules and Accuracy Rate between the Yuan-and-Shaw's Method [19] and the Proposed Method

	Yuan-and-Shaw's Method [19]	The Proposed Method (under $\alpha = 0.9$ and $\beta = 0.6$)

Number of Rules	6	3
Accuracy Rate	81.25%	93.75%

5. Conclusions

In this paper, we have presented a new method for generating fuzzy rules from numerical data for handlin fuzzy classification problems based on the fuzzy subsethood values between the decisions to be made and terms of attributes of subgroups by using the level threshold value α and the applicability threshold value β, where $\alpha \in [0,1]$ and $\beta \in [0,1]$. We apply the proposed method to deal with the Saturday Morning Problem [19]. The proposed method is better than the one presented in [19] due to the fact that

(1) The proposed method gets a better accuracy rate than the one presented in [19]. From the experimental results, we can see that the accuracy rate of the proposed method is 93.75% (under $\alpha = 0.9$ and $\beta = 0.6$), while the accuracy rate of the Yuan-and-Shaw's method is 81.25%.
(2) The proposed method generates fewer fuzzy rules than the one presented in [19]. From the experime ntal results, we can see that the number of fuzzy rules generate by the proposed method is 3, but the number of fuzzy rules generated by Yuan-and-Shaw's method is 6.
(3) The proposed method needs less calculations than the one presented in [19].

Acknowledgements

This work was supported by the National Science Council, Republic of China, under Grant NS 89-2213-E-011-060.

References

[1] J. L. Castro and J. M. Zurita, "An inductive learnin algorithm in fuzzy systems," *Fuzzy Sets and Systems*, vol. 89, no. 2, pp. 193-203, 1997.
[2] R. L. P. Chang and T. Pavlidis, "Fuzzy decision tree algorithms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-7, no. 1, pp. 28-35, 1977.
[3] S. M. Chen and M. S. Yeh, "Generating fuzzy rules from relational database systems for estimating null values," *Cybernetics and Systems: An International Journal*, vol. 29, no. 6, pp. 363 – 376, 1998.
[4] J. Giarratano and G. Riley, *Expert Systems: Principles and Programming*. Boston: PWS, 1994.
[5] Y. Hayashi and A. Imura, "Fuzzy neural expert system with automated extraction of fuzzy if-then rules from a trained neural network," in *Proc. 1990 First Inter. Symp. on Uncertainty Modeling and Analysis*, University of Maryland, College Park, Maryland, U. S. A., 1990.
[6] A. Homaifar and E. McCormick, "Simultaneous design of membership functions and rule sets for fuzz controllers using genetic algorithms," *IEE Transactions on Fuzzy Systems*, vol. 3, no. 2, pp 129-139, 1995.
[7] T. P. Hong and C. Y. Lee, "Induction of fuzzy rules and membership functions from training examples," *Fuzz Sets and Systems*, vol. 84, no. 1, pp. 33-47, 1996.
[8] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671 – 680, 1983.
[9] F. Klawonn and R. Kruse, "Constructing a fuzzy controller from data," *Fuzzy Sets and Systems*, vol. 85 no. 2, pp. 177-193, 1997.
[10] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Englewood Cliffs: Prentice-Hall, 1995.
[11] T. Kohonen, *Self-Organization and Associative Memory*. Berlin: Springer, 1988.
[12] B. Kosko, "Fuzzy entropy and conditioning," *Information Sciences*, vol. 40, no. 2, pp. 165 -174, 1986.
[13] C. C. Lee, "Fuzzy logic in control systems: fuzzy logic controller- part I," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp. 404 -418, 1990.
[14] K. Nozaki, H. Ishibuchi, and H. Tanaka, "A simple but powerful heuristic method for generating fuzzy rules from numerical data," *Fuzzy Sets and Systems*, vol. 86 no. 3, pp. 251-270, 1997.
[15] J. R. Quinlan, "Simplifying decision trees," *International Journal of Man-Machine Studies*, vol. 27, no. 3, pp. 221-234, 1987.
[16] J. R. Quinlan, "Decision trees and decision making," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp.339-346, 1994.
[17] T. P. Wu and S. M. Chen, "A new method for constructing membership functions and fuzzy rules from training examples," *IEEE Transactions on Systems, Man, and Cybernetics - Part B*, vol. 29, no. 1, pp. 25-40, 1999.
[18] Y. Yoshinari, W. Pedrycz, and K. Hirota, "Construction of fuzzy models through clustering techniques," *Fuzzy Sets and Systems*, vol. 54, no. 2, pp. 157-165, 1993.
[19] Y. Yuan and M. J. Shaw, "Induction of fuzzy decision trees," *Fuzzy Sets an d Systems*, vol. 69, no. 2, pp. 125-139, 1995.
[20] Y. Yuan and H. Zhuang, "A genetic algorithm for generating fuzzy classification rules," *Fuzzy Sets and Systems*, vol. 84, no. 1, pp. 1-19, 1996.
[21] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338-353, 1965.
[22] L. A. Zadeh, "Fuzzy logic," *IEEE Computer*, vol. 21, no. 4, pp. 83-91, 1988.
[23] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning - I," *Information Sciences*, vol. 8, no. 3, pp. 199-249, 1975.
[24] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications*. Boston: Kluwer-Nijhoff, 1991.