A NEW METHOD TO GENERATE FUZZY RULES FROM NUMERICAL DATA BASED ON THE EXCLUSION OF ATTRIBUTE TERMS

Chi-Hao Chang and Shyi-Ming Chen

Department of Electronic Engineering
National Taiwan University of Science and Technology
Taipei, Taiwan, R. O. C.

ABSTRACT

To develop a fuzzy system, the most important task is to derive a set of fuzzy rules from a set of training data. In recent years, many methods have been developed to automatically derive fuzzy rules from training instances. In this paper, we present a new method to generate fuzzy rules from numerical data based on the exclusion of attribute terms to deal with the Iris data classification problem. The experimental results show that the proposed method can get a higher classification accuracy rate than the existing methods.

1. INTRODUCTION

In the real-world, many things contain uncertainty and vagueness. Fuzzy set theory [21] has been developed to deal with uncertainty and vagueness [21]. To develop a fuzzy system, the most important task is to find a set of fuzzy rules. This can be done by two approaches. One is by acquiring the knowledge from domain experts through knowledge acquisition tools. But the domain experts may not be available and the process of knowledge acquisition may be very time consuming. The other approach is by applying machine learning methods, such that the system can automatically generate fuzzy IF-THEN rules from a set of training instances. In recent years, fuzzy systems that can automatically derive fuzzy rules from training instances have been developed [2]-[5], [7]-[10], [14], [16], [18], [19].

In this paper, we present a new method for fuzzy rules generation based on the exclusion of attribute terms, i.e., instead of choosing only one attribute term whose fuzzy subsethood value exceeds the level threshold value α for each attribute, we exclude the attribute terms whose complement of fuzzy subsethood values exceed the level threshold value α, where α ∈ [0, 1]. We also apply the proposed method to deal with the Iris data classification problem. The experimental results show that the proposed method for generating fuzzy rules from numerical data has a higher classification accuracy rate than the existing methods.

The rest of this paper is organized as follows. In Section 2, we briefly review the concepts of fuzzy sets from [12], [20], and [21]. In Section 3, we present a new algorithm for fuzzy rules generation from training data. In section 4, we use a simple example to illustrate the proposed algorithm. In section 5, we show the experimental results of the proposed algorithm to deal with the classification problem. The conclusions are discussed in Section 6.

2. BASIC CONCEPTS OF FUZZY SETS

In [21], Zadeh proposed the theory of fuzzy sets. Let U be the universe of discourse, $U = \{u_1, u_2, \ldots, u_n\}$. A fuzzy set A of the universe of discourse U can be characterized by a membership function μ_A, μ_A: U → [0, 1], represented by

$$A = \mu_A(u_1)/u_1 + \mu_A(u_2)/u_2 + \mu_A(u_3)/u_3 + \ldots + \mu_A(u_n)/u_n,$$

where $\mu_A(u_i)$ indicates the grade of membership of u_i in the fuzzy set A and $\mu_A(u_i) \in [0, 1]$. In [5], we have presented a method to generate fuzzy IF-THEN rules from numerical data based on the fuzzy subsethood measure S [12], [20]. The definition of the fuzzy subsethood measure S is as follows.

Definition 2.1: Let A and B be two fuzzy sets defined in a finite universe of discourse U with membership functions μ_A and μ_B, respectively. The fuzzy subsethood S(A, B) measures the degree in which A is a subset of B shown as follows:

$$S(A,B) = \frac{M(A \cap B)}{M(A)} = \frac{\sum_{u \in U} \min(\mu_A(u), \mu_B(u))}{\sum_{u \in U} \mu_A(u)} \tag{1}$$

where S(A, B) ∈ [0, 1].

Definition 2.2: Let A and B be two fuzzy sets of the universe of discourse U. If A is a subset of B, then $\mu_A(u) \le \mu_B(u)$, $\forall u \in U$.

3. A NEW ALGORITHM FOR FUZZY RULES GENERATION

In this section, we present a new method to generate fuzzy rules from numerical data. According to the fuzzy subsethood measure, the method presented in [5] selects only one attribute term for each attribute. In this paper, we don't select only one attribute term for each attribute but exclude all the attribute terms whose complement of fuzzy subsethood measure exceed the level threshold value α, where $\alpha \in [0,1]$. The complement of fuzzy subsethood measure is defined as follows.

Definition 3.1: Let A and B be two fuzzy sets defined in a finite universe of discourse U. The fuzzy subsethood in which A is a subset of B is denoted as S(A, B), where $S(A,B) = \dfrac{M(A \cap B)}{M(A)}$ and $S(A, B) \in [0, 1]$. Then, the complement $\overline{S(A,B)}$ of the fuzzy subsethood S(A, B), is defined by:

$$\overline{S(A,B)} = 1 - S(A,B), \qquad (2)$$

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

For simplicity, we use a quadruple (a_1, a_2, a_3, a_4) to represent an attribute term, where the definition of an attribute term is defined as follows.

Definition 3.2: In a fuzzy classification problem, the values of an attribute are linguistic terms called attribute terms, where the attribute terms are represented by fuzzy sets.

There are four kinds of membership functions which can be used to represent an attribute term, i.e., the left-trapezoidal membership function, the full-trapezoidal membership function, the right-trapezoidal membership function, and the triangular membership function, where the four kinds of membership functions are defined as follows:

(1) Left-trapezoidal membership function:

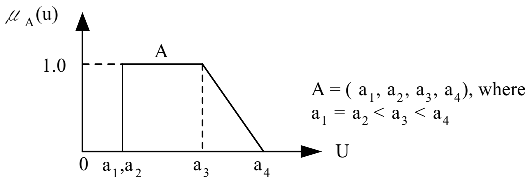

Fig. 1. Left-trapezoidal membership function.

(2) Full-trapezoidal membership function:

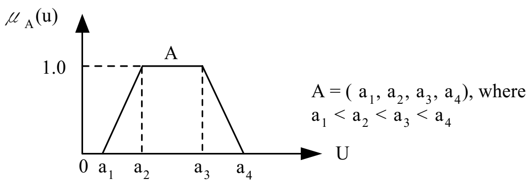

Fig. 2. Full-trapezoidal membership function.

(3) Right-trapezoidal membership function:

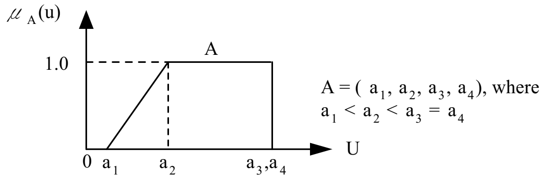

Fig. 3. Right-trapezoidal membership function.

(4) Triangular membership function:

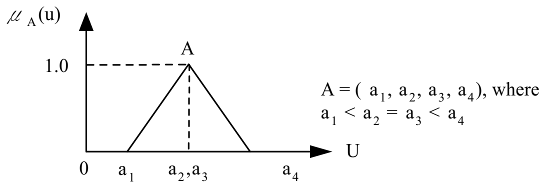

Fig. 4. Triangular membership function.

Definition 3.3: Let A and B be two fuzzy sets defined in a finite universe of discourse U with the membership functions μ_A and μ_B, respectively. If $\bigvee_{u \in U} (\mu_A(u) \wedge \mu_B(u)) > 0$, then A and B are called adjacent, where "$\vee$" and "$\wedge$" are the maximum operator and the minimum operator, respectively.

In the following, we present a procedure to merge attribute terms of an attribute, where attribute terms are represented by fuzzy sets. Let T be a set of attribute terms of an attribute. The procedure to merge the attribute terms in the set T is presented as follows:

```
Procedure Merge(T)
    WHILE T is not empty DO
            Select an attribute term A from T
            IF there exists another term B in T such that A
               and B are adjacent
            THEN Call Procedure Simplify(A, B) to form
                    a new term C and add C into T, and
                    then delete A and B from T
            ELSE delete A from T
    END
END.
```

In the procedure **Merge**(T), we can see that the procedure **Simplify**(A, B) is to merge two attribute terms into a new attribute term. Suppose two attribute terms A = (a_1, a_2, a_3, a_4) and B = (b_1, b_2, b_3, b_4) can be merged to form a new attribute term C = (c_1, c_2, c_3, c_4). The procedure **Simplify**(A, B) is presented as follows:

```
Procedure Simplify(A, B) /* A = (a1, a2, a3, a4) and B = (b1, b2, b3, b4) */
    IF a1 ≥ b1
    THEN c1 = a1, c2 = a2, c3 = b3, c4 = b4
    ELSE c1 = b1, c2 = b2, c3 = a3, c4 = a4
    RETURN C    /* C = (c1, c2, c3, c4) */
END.
```

In [11], Kao has presented a method to calculate the degrees of fuzziness of the attributes. The method is reviewed as follows. Assume that there are *n* training instances in the training data set and assume that A is an attribute of the training data set. Assume that the maximum value of the attribute A is denoted by Max(A), the minimum value of the attribute A is denoted by Min(A), and the average value of the attribute A is denoted by Ave(A). Let X_i be the ith value of the attribute A, where $1 \le i \le n$. Then, the degree of fuzziness D(A) of the attribute A is calculated as follows:

$$D(A) = \frac{\sum_{i=1}^{n} D_i}{n}, \quad (3)$$

where

$$D_i = \begin{cases} \dfrac{X_i - \text{Ave}(A)}{\text{Max}(A) - \text{Ave}(A)}, & \text{If } X_i \ge \text{Ave}(A) \\ \dfrac{\text{Ave}(A) - X_i}{\text{Ave}(A) - \text{Min}(A)}, & \text{If } X_i < \text{Ave}(A) \end{cases} \quad (4)$$

$D_i \in [0, 1]$, $1 \le i \le n$, and $D(A) \in [0, 1]$. The larger the value D(A), the more helpful the attribute A is to be used for classification. After calculating the degree of fuzziness of each attribute, if the degree of fuzziness D(A) of an attribute A is smaller than a threshold value λ given by the user, where $\lambda \in [0, 1]$, then the attribute A is regarded as a useless attribute for classification and it is deleted form the training data set.

The proposed algorithm for fuzzy rules generation from numerical data is presented as follows:

Step 1: Remove the useless attributes from the training data set based on formulas (3) and (4).

Step 2: Fuzzify each training instance.

Step 3: Divide the training instances into several groups according to the decisions of the training instances, i.e., training instances with the same decision will be grouped together.

Step 4: According to formula (1), for each group derived in Step 3, calculate the fuzzy subsethood values between the decision to be made and each attribute term of each attribute.

Step 5: According to formula (2), calculate the complement of the fuzzy subsethood values derived in Step 4.

Step 6: Let T be a set containing attribute terms.

FOR each group derived in Step 2 **DO**
  FOR each attribute in the group **DO**
    Let T be an empty set.
    FOR each attribute term of the attribute **DO**
      IF the complement of the fuzzy subsethood value between the attribute term and the decision to be made does not exceed the level threshold value α given by the user, where $\alpha \in [0, 1]$
      THEN add the attribute term into T,
    END
    Call Procedure **Merge**(T) to merge the attribute terms in the set T
  END
END.

Step 7: Generate fuzzy IF-THEN rules from the generated membership functions of the generated attribute terms derived in Step 6. For example, in the Iris data classification problem, suppose the generated membership functions of the attributes “Sepal Length”, “Sepal Width”, “Petal Length”, and “Petal Width” for the group “Setosa” derived in Step 6 are “Setosa(SL)”, “Setosa(SW)”, “Setosa(PL)”, and “Setosa(PW)”, respectively. Then, the fuzzy IF-THEN rule generated for the group “Setosa” is:

IF SL is Setosa(SL) and SW is Setosa(SW) and PL is Setosa(PL) and PW is Setosa(PW)
THEN the flower is Setosa.

4. AN EXAMPLE

In this section, we use a simple example to illustrate the proposed algorithm. For simplicity, we only chose 15 instances randomly from the Iris data [6] as a simple example to illustrate the proposed algorithm. There are three species of flower in the Iris data, i.e., “Setosa”, “Versicolor”, and “Verginica”, and there are 150 instances in the Iris data, with 50 instances for each species, and each species with four attributes, i.e., Sepal Length (SL), Sepal Width (SW), Petal Length (PL), and Petal Width (PW). In this paper, the membership functions for each attribute which we used to fuzzify the training instances are adopted from [3] as shown in Fig. 5 to Fig. 8.

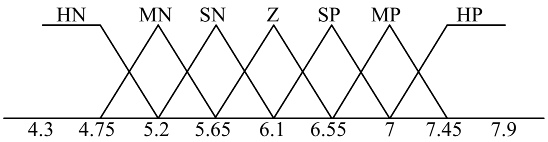

Fig. 5. Membership function for the attribute Sepal Length.

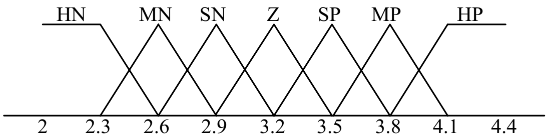

Fig. 6. Membership function for the attribute Sepal Width.

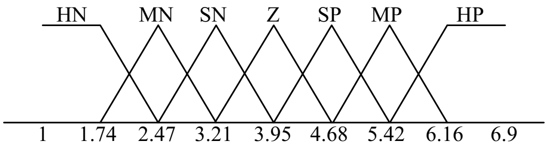

Fig. 7. Membership function for the attribute Petal Length.

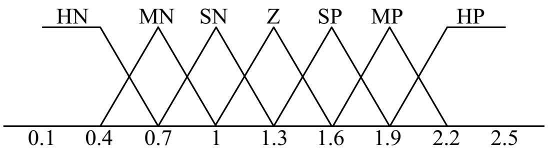

Fig. 8. Membership function for the attribute Petal Width.

In order to clearly illustrate the fuzzy rules generation process of the proposed algorithm, we only chose 15 instances from the Iris data for illustration, where 5 instances for each species (i.e., Setosa, Versicolor, and Verginica) are chosen. The chosen instances for this example are shown in Table 1.

Table 1. A Subset of the Iris Data

Attributes / Training Instances	SL	SW	PL	PW	Species
1	5	3.3	1.4	0.2	Setosa
2	5.7	2.8	4.1	1.3	Versicolor
3	5.9	3	5.1	1.8	Verginica
4	5	3.6	1.4	0.2	Setosa
5	6.5	2.8	4.6	1.5	Versicolor
6	6.5	3	5.8	2.2	Verginica
7	4.8	3	1.4	0.1	Setosa
8	6	2.2	4	1	Versicolor
9	6.8	3	5.5	2.1	Verginica
10	5.7	3.8	1.7	0.3	Setosa
11	6.2	2.2	4.5	1.5	Versicolor
12	7.7	2.6	6.9	2.3	Verginica
13	5.1	3.4	1.5	0.2	Setosa
14	5.5	2.5	4	1.3	Versicolor
15	6.9	3.1	5.4	2.1	Verginica

The step-by-step demonstration of the proposed algorithm is shown as follows:

[Step 1] /* Remove the useless attributes from the training data set based on formulas (3) and (4) */

Suppose the user set the level threshold value $\lambda = 0.5$. Then, based on formula (3) and (4), the degrees of fuzziness of the attributes SL, SW, PL and PW shown in Table 1 can be calculated and the results are D(SL) = 0.47, D(SW) = 0.44, D(PL) = 0.58, and D(PW) = 0.64. Because the degrees of fuzziness of the attributes SL and SW are less than the threshold value λ, where $\lambda = 0.5$, we can see that SL and SW are useless attributes for classification, thus the attributes SL and SW are removed from the training data set. Thus, the resulting training data set is shown in Table 2.

Table 2. A Subset of the Iris Data After Removing the Attributes SL and SW

Attributes / Training Instances	PL	PW	Species
1	1.4	0.2	Setosa
2	4.1	1.3	Versicolor
3	5.1	1.8	Verginica
4	1.4	0.2	Setosa
5	4.6	1.5	Versicolor
6	5.8	2.2	Verginica
7	1.4	0.1	Setosa
8	4	1	Versicolor
9	5.5	2.1	Verginica
10	1.7	0.3	Setosa
11	4.5	1.5	Versicolor
12	6.9	2.3	Verginica
13	1.5	0.2	Setosa
14	4	1.3	Versicolor
15	5.4	2.1	Verginica

[Step 2] /* Fuzzify each training instance */

According to the membership functions shown in Fig. 5 to Fig. 8, the training instances shown in Table 2 can be fuzzified as shown in Table 3.

[Step 3] /* Divide the training instances into several groups according to the decisions of the training instances, i.e., training instances with the same decision will be grouped together */

From Table 3, we can see that training instances 1, 4, 7, 10, and 13 have the same decision (i.e., Setosa), so they can be grouped together denoted as "Group(Setosa)". Similarly, training instances 2, 5, 8, 11, and 14 have the same decision (i.e., Versicolor), so they can be grouped together denoted as "Group(Versicolor)", and training instances 3, 6, 9, 12, and 15 have the same decision (i.e., Verginica), so they can be grouped together denoted as "Group(Verginica)". The results are shown in Table 4.

[Step 4] /* According to formula (1), for each group derived in Step 3, calculate the fuzzy subsethood values between the decision to be made (i.e., species) and each attribute term */

According to formula (1), the fuzzy subsethood values between different species of flower and the attribute terms of the attributes "Petal Length" and "Petal Width" can be calculated, respectively. The results of the calculations are shown in Table 5.

[Step 5] /* According to formula (2), calculate the complement of the fuzzy subsethood values derived in Step 4 */

According to formula (2), the complement of the fuzzy subsethood values shown in Table 5 are shown in Table 6.

[Step 6] /* Merge membership functions of each attribute */

For example, suppose the user set the level threshold value α = 0.95, and we consider the attribute PL for the species "Setosa". From Table 6, we can see that the complement of the fuzzy subsethood values between the decision to be made (i.e., Setosa) and the attribute terms MN, SN, Z, SP, MP and HP exceed the level threshold value α, where α = 0.95, so these six attribute terms will be excluded from the set T of attribute terms, i.e., only the attribute term HN will be added into the set T. After calling the procedure **Merge**(T), the initial membership function for the attribute PL shown in Fig. 7 can be merged into the membership function "Setosa(PL)" as shown in Fig. 9.

Similarly, consider the attribute PW for the species "Setosa". From Table 6, we can see that the complement of the fuzzy subsethood values between the decision to be made (i.e., Setosa) and the attribute terms MN, SN, Z, SP, MP and HP exceed the level threshold value α, where α = 0.95, so these six attribute terms will be excluded from the set T of attribute terms, i.e., only the attribute term HN will be added into the set T. After calling the procedure **Merge**(T), the initial membership function for the attribute PW shown in Fig. 8 can be merged into the membership function "Setosa(PW)" as shown in Fig. 10.

By the same way, the initial membership functions for the attributes PL and PW of species Versicolor shown in Fig. 7 and Fig. 8, respectively can be merged into the membership functions "Versicolor(PL)" and "Versicolor(PW)" as shown in Fig. 11 and Fig. 12, respectively. The initial membership functions for the attributes PL and PW of species Verginica shown in Fig. 7 and Fig. 8, respectively can be merged into the membership functions "Verginica(PL)" and "Verginica(PW)", as shown in Fig. 13 and Fig 14, respectively.

Table 3. The Fuzzified Training Instances

Attributes / Training Instances	PL							PW							Species		
	HN	MN	SN	Z	SP	MP	HP	HN	MN	SN	Z	SP	MP	HP	Setosa	Versicolor	Verginica
1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
2	0	0	0	0.79	0.21	0	0	0	0	0	1	0	0	0	0	1	0
3	0	0	0	0	0.43	0.57	0	0	0	0	0	0.33	0.67	0	0	0	1
4	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
5	0	0	0	0.11	0.89	0	0	0	0	0	0.33	0.67	0	0	0	1	0
6	0	0	0	0	0	0.49	0.51	0	0	0	0	0	0	1	0	0	1
7	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
8	0	0	0	0.93	0.07	0	0	0	0	1	0	0	0	0	0	1	0
9	0	0	0	0	0	0.89	0.11	0	0	0	0	0	0.33	0.67	0	0	1
10	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
11	0	0	0	0.25	0.75	0	0	0	0	0	0.33	0.67	0	0	0	1	0
12	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1
13	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
14	0	0	0	0.93	0.07	0	0	0	0	0	1	0	0	0	0	1	0
15	0	0	0	0	0.03	0.97	0	0	0	0	0	0	0.33	0.67	0	0	1

Table 4. Group Together the Training Instances Shown in Table 3

Training Instances \ Attributes	PL							PW							Species		
	HN	MN	SN	Z	SP	MP	HP	HN	MN	SN	Z	SP	MP	HP	Setosa	Versicolor	Verginica
1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
4	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
7	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
10	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
13	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
2	0	0	0	0.79	0.21	0	0	0	0	0	1	0	0	0	0	1	0
5	0	0	0	0.11	0.89	0	0	0	0	0	0.33	0.67	0	0	0	1	0
8	0	0	0	0.93	0.07	0	0	0	0	1	0	0	0	0	0	1	0
11	0	0	0	0.25	0.75	0	0	0	0	0	0.33	0.67	0	0	0	1	0
14	0	0	0	0.93	0.07	0	0	0	0	0	1	0	0	0	0	1	0
3	0	0	0	0	0.43	0.57	0	0	0	0	0	0.33	0.67	0	0	0	1
6	0	0	0	0	0	0.49	0.51	0	0	0	0	0	0	1	0	0	1
9	0	0	0	0	0	0.89	0.11	0	0	0	0	0	0.33	0.67	0	0	1
12	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1
15	0	0	0	0	0.03	0.97	0	0	0	0	0	0	0.33	0.67	0	0	1

Table 5. The Fuzzy Subsethood Values Between Attribute Terms and Species

Species \ Attributes	PL							PW						
	HN	MN	SN	Z	SP	MP	HP	HN	MN	SN	Z	SP	MP	HP
Setosa	1	0	0	0	0	0	0	1	0	0	0	0	0	0
Versicolor	0	0	0	0.60	0.40	0	0	0	0	0.20	0.53	0.27	0	0
Verginica	0	0	0	0	0.09	0.58	0.32	0	0	0	0	0.07	0.27	0.67

Table 6. The Complement of Fuzzy Subsethood Values between Attribute Terms and Species

Species \ Attributes	PL							PW						
	HN	MN	SN	Z	SP	MP	HP	HN	MN	SN	Z	SP	MP	HP
Setosa	0	1	1	1	1	1	1	0	1	1	1	1	1	1
Versicolor	1	1	1	0.40	0.60	1	1	1	1	0.80	0.47	0.73	1	1
Verginica	1	1	1	1	0.91	0.42	0.68	1	1	1	1	0.93	0.73	0.33

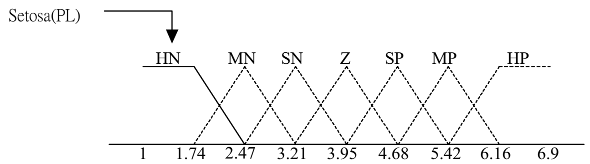

Fig. 9. Membership function Setosa(PL) for the attribute PL of the species Setosa.

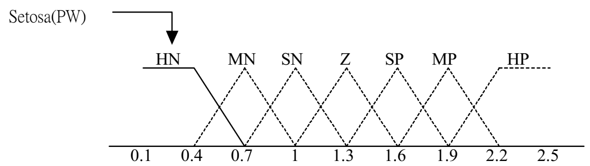

Fig. 10. Membership function Setosa(PW) for the attribute PW of the species Setosa.

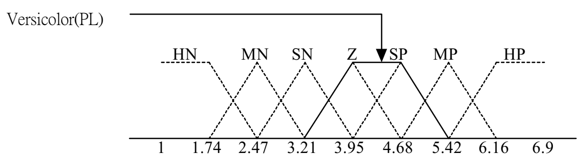

Fig. 11. Membership function Versicolor(PL) for the attribute PL of the species Versicolor.

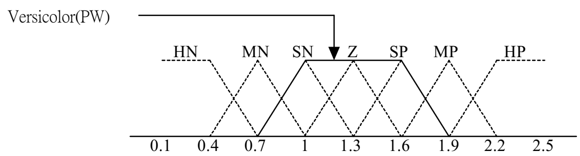

Fig. 12. Membership function Versicolor(PW) for the attribute PW of the species Versicolor.

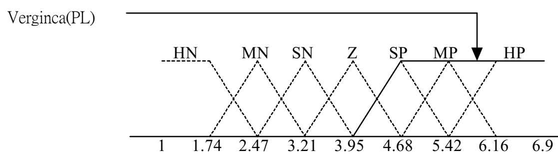

Fig. 13. Membership function Verginica(PL) for the attribute PL of the species Verginica.

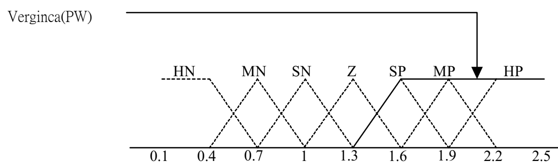

Fig. 14. Membership function Verginica(PW) for the attribute PW of the species Verginica.

[Step 7] /* Generate fuzzy IF-THEN rules from the simplified membership functions derived in Step 6 */

From Fig. 9 to Fig. 14, the fuzzy IF-THEN rules generated from the 15 training instances by the proposed algorithm are shown as follows:

Rule 1: IF PL is Setosa(PL) and PW is Setosa(PW)
THEN the flower is Setosa

Rule 2: IF PL is Versicolor(PL) and PW is Versicolor(PW)
THEN the flower is Versicolor

Rule 3: IF PL is Verginica(PL) and PW is Verginica(PW)
THEN the flower is Verginica

5. EXPERIMENTAL RESULTS

We chose 120 instances randomly among the 150 instances of the Iris data as the set of training instances and the remaining 30 instances of the Iris data as the set of testing instances. The experimental results are shown in Table 7, where different numbers of attribute terms for the attributes are considered (i.e., 5, 7, 9, and 11), each test was executed 200 runs, and the level threshold value α is 0.95.

Table 7. Experimental Results (with 120 Instances as Training Instances and the Remaining 30 Instances as Testing Instances) for Different Numbers of Attribute Terms for the Attributes

Classification Accuracy Rate / Test (Each Was Executed 200 Runs) \ Number of Attribute Terms	5	7	9	11
1	0.9538	0.9728	0.9543	0.9590
2	0.9518	0.9753	0.9553	0.9582
3	0.9568	0.9710	0.9583	0.9583
4	0.9577	0.9688	0.9545	0.9618
5	0.9533	0.9712	0.9587	0.9542
6	0.9537	0.9733	0.9542	0.9620
7	0.9550	0.9702	0.9563	0.9592
8	0.9540	0.9727	0.9563	0.9570
9	0.9547	0.9695	0.9520	0.9627
10	0.9523	0.9743	0.9607	0.9578

We also chose 75 instances randomly among the Iris data as the set of training instances, and the other 75 instances of the Iris data as the set of testing instances. The experimental results are shown in Table 8, where different numbers of attribute terms for the attributes are considered (i.e., 5, 7, 9, and 11), each test was executed 200 runs, and the level threshold value α is 0.95.

Table 8. Experimental Results (with 75 Instances as Training Instances and the Other 75 Instances as Testing Instances) for Different Numbers of Attribute Terms for the Attributes

Classification Accuracy Rate / Test (Each Was Executed 200 Runs) \ Number of Attribute Terms	5	7	9	11
1	0.9557	0.9671	0.9621	0.9553
2	0.9558	0.9682	0.9552	0.9477
3	0.9535	0.9663	0.9565	0.9466
4	0.9529	0.9662	0.9527	0.9581
5	0.9537	0.9683	0.9522	0.9511
6	0.9531	0.9672	0.9524	0.9509
7	0.9499	0.9663	0.9551	0.9503
8	0.9523	0.9664	0.9585	0.9497
9	0.9539	0.9649	0.9547	0.9523
10	0.9542	0.9663	0.9544	0.9400

From Table 7 and Table 8, we can see that different numbers of attribute terms for the attributes can affect the classification accuracy rate. When the number of attribute terms for each attribute is equal to seven, we can get the highest classification accuracy rate. With 75 instances of the Iris data as the training data set and the remaining 75 instances of the Iris data as the testing data set, and when the number of attribute terms for each attribute is 7, the average classification accuracy rate of the proposed method after 2000 runs is 96.67%. With 120 instances of the Iris data as the training data set and the remaining 30 instances of the Iris data as the testing data set, and when the number of attribute terms for each attribute is 7, the average classification accuracy rate of the proposed method after 2000 runs is 97.19%.

In [7], Hong and Lee presented a general learning method to generate membership functions and fuzzy IF-THEN rules from a set of given training examples. The average classification accuracy rate of the Hong-and-Lee's method presented in [7] after 200 runs is 95.57% and the average number of fuzzy IF-THEN rules generated by the method presented in [7] is 6.21. In [18], Wu and Chen presented a method to construct membership functions and fuzzy IF-THEN rules from training examples based on the α-cuts of equivalence relations and the α-cuts of fuzzy sets. For the Iris data classification problem, the average accuracy rate of the method presented in [18] after 200 runs is 96.21% and the number of fuzzy IF-THEN rules generated by the method presented in [18] is 3, where both Hong-and-Lee's method and Wu-and-Chen's method chose 75 instances randomly as the training data set and the other 75 instances of the Iris data as the testing data set.

6. CONCLUSIONS

In this paper, we have presented a new method for fuzzy rules generation to deal with the Iris data classification problem. We aslo compared the proposed method with Hong-and-Lee's method presented in [7] and Wu-and-Chen's method presented in [18]. From the experimental results shown in Section 5, we can see that the proposed method have the following advantages:

(1) The average classification accuracy rate of the proposed method is higher than Hong-and-Lee's method presented in [7], and Wu-and-Chen's method presented in [18] when the number of attribute terms for each attribute is 7.

(2) The proposed method generates fewer fuzzy rules than Hong-and-Lee's method presented in [7].

ACKNOWLEDGEMENTS

This work was supported in part by the National Science Council, Republic of China, under Grant NSC 89-2213-E-011-100.

REFERENCES

[1] B. G. Buchanan and E. H. Shortliffe, *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Projects*. MA: Addison-Wesley, 1984.

[2] D. G. Burkhardt and P. P. Bonissone, "Automated fuzzy knowledge base generation and tuning," *Proceedings of the 1992 IEEE International Conference on Fuzzy Systems*, San Diego, California, pp. 179-188, 1992.

[3] J. L. Castro, J. J. Castro-Schez, and J. M. Zurita, "Learning maximal structure rules in fuzzy logic for knowledge acquisition in expert systems," *Fuzzy Sets and Systems*, vol. 101, no. 3, pp. 331-342, 1999.

[4] 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. 28, no. 8, pp. 695-723, 1999.

[5] S. M. Chen, S. H. Lee, and C. H. Lee, "Generating fuzzy rules from numerical data for handling fuzzy classification problems," *Proceedings of the 1999 National Computer Symposium*, Taipei, Taiwan, Repulic of China, vol. 2, pp. 336-343, 1999.

[6] R. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics Part 2*, vol. 7, pp. 179-188, 1936.

[7] T. P. Hong and C. Y. Lee, "Induction of fuzzy rules and membership functions from training examples," *Fuzzy Sets and Systems*, vol. 84, no. 1, pp. 33-47, 1996.

[8] T. P. Hong and C. Y. Lee, "Effect of merging order on performance of fuzzy induction," *Intelligent Data Analysis*, vol. 3, no. 2, pp. 139-151, 1999.

[9] T. P. Hong and J. B. Chen, "Finding relevant attributes and membership functions," *Fuzzy Sets and Systems*, vol. 103, no. 3, pp. 389-404, 1999.

[10] T. P. Hong and J. B. Chen, "Processing individual fuzzy attributes for fuzzy rule induction," *Fuzzy Sets and Systems*, vol. 112, no. 1, pp. 127-140, 2000.

[11] C. H. Kao, "A new method to generate fuzzy rules from training data containing noise for handling classification problems," *Master Thesis*, Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R. O. C., 2000.

[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 system: fuzzy logic controller – Parts I and II," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp. 404-435, 1990.

[14] H. Nomura, I. Hayashi, and N. Wakami, "A learning method of fuzzy inference rules by descent method," *Proceedings of the 1992 IEEE International Conference on Fuzzy Systems*, San Diego, California, pp. 203-210, 1992.

[15] G. Shafer and R. Logan, "Implementing Dempster's rule for hierarchical evidence," *Artificial Intelligence*, vol. 33, no. 3, pp. 271-298, 1987.

[16] T. Sudkamp and R. J. Hammell II, "Interpolation, completion, and learning fuzzy rules," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 24, no. 2, pp. 332-342, 1994.

[17] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, no. 1, pp. 116-132, 1985.

[18] 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*, vol. 29, no. 1, pp. 25-40, 1999.

[19] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 6, pp. 1414-1427, 1992.

[20] Y. Yuan and M. J. Shaw, "Induction of fuzzy decision trees," *Fuzzy Sets and Systems*, vol. 69, no. 2, pp. 125-139, 1995.

[21] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338-353, 1965.

[22] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications*. Boston: Kluwer Academic Publisher, 1991.