A New Method for Constructing Membership Functions and Fuzzy Rules from Training Examples

Tzu-Ping Wu and Shyi-Ming Chen, *Senior Member, IEEE*

***Abstract*—To extract knowledge from a set of numerical data and build up a rule-based system is an important research topic in knowledge acquisition and expert systems. In recent years, many fuzzy systems that automatically generate fuzzy rules from numerical data have been proposed. In this paper, we propose a new fuzzy learning algorithm based on the α-cuts of equivalence relations and the α-cuts of fuzzy sets to construct the membership functions of the input variables and the output variables of fuzzy rules and to induce the fuzzy rules from the numerical training data set. Based on the proposed fuzzy learning algorithm, we also implemented a program on a Pentium PC using the MATLAB development tool to deal with the Iris data classification problem. The experimental results show that the proposed fuzzy learning algorithm has a higher average classification ratio and can generate fewer rules than the existing algorithm.**

***Index Terms*—Fuzzy learning algorithms, fuzzy rules, knowledge acquisition, membership functions, rule-based systems.**

I. Introduction

In RECENT years, expert systems have become more and more popular and important in many applications. The purpose of expert systems is to emulate the reasoning process of human experts within a specific domain of knowledge. An expert system consists of a knowledge base, a user interface, and an inference engine [8]. Knowledge engineering plays an important role in the research field of expert systems. Knowledge engineering involves knowledge acquisition, knowledge representation, and human–machine interaction. The purpose of knowledge acquisition is to extract knowledge from the opinions of experts or a set of numerical data. It is difficult to build up a conventional mathematical model to elicit knowledge from experts or the sample data in the real-world application due to the fact that it needs to precisely describe all the characteristics of the system and it lacks flexibility. Fuzzy set theory proposed by Zadeh [38] can deal with the vagueness and uncertainty residing in the knowledge possessed by human beings or implicated in the numerical data, and it allows us to represent the system parameters with linguistic terms. Fuzzy rules have been used as a key tool to express knowledge in fuzzy logic and are more adequate and flexible than the traditional IF–THEN rules. In [6], Dubois *et al.* have surveyed different possible semantics for fuzzy rules and showed how they can be captured in the framework of fuzzy sets and possibility theory.

There are different approaches to extract knowledge from experts or training examples. These methods are based on neural networks or the fuzzy set theory [9]–[13], [15], [16], [20], [21], [26], [29], [32]. In [32], Wang *et al.* used a table-lookup scheme to generate fuzzy rules directly from numerical examples and proved that a fuzzy inference system is a universal approximator by the Stone–Weierstrass theorem [33]. In [26], Nozaki *et al.* presented a heuristic method for generating Takagi–Sugeno–Kang (TSK) fuzzy rules from numerical data, and then translated the consequent parts of TSK fuzzy rules into linguistic representation. In [9], Grauel *et al.* have investigated the connection between the shape of transfer functions and the shape of membership functions, where membership functions for multi-input of Sugeno controllers and designed rules were derived. In [16], Klawonn *et al.* discussed how fuzzy clustering techniques could be applied to construct a fuzzy controller from the training data. In [11], Hong and Lee have pointed out that the drawbacks of most fuzzy controllers and fuzzy expert systems are that they need to predefine membership functions and fuzzy rules to map numerical data into linguistic terms and to make fuzzy reasoning work. They proposed a method based on the fuzzy clustering technique and the decision tables to derive membership functions and fuzzy rules from numerical data. However, Hong and Lee's algorithm presented in [11] needs to predefine the membership functions of the input linguistic variables and it simplifies fuzzy rules by a series of merge operations. As the number of variables becomes larger, the decision table will grow tremendously and the process of the rule simplification based on the decision tables becomes more complicated. Thus, we must develop a new algorithm to overcome the drawbacks of the Hong and Lee's algorithm.

In this paper, we propose a new fuzzy learning algorithm based on the α-cuts of fuzzy equivalence relations and the α-cuts of fuzzy sets to divide numerical data into different partitions and to automatically derive membership functions for each partition. Based on the hierarchical relationships between different data partitions, the membership functions of the input linguistic variables and the output linguistic variables can be constructed automatically and the fuzzy rules can be generated directly.

The rest of this paper is organized as follows. In Section II, basic concepts and definitions of the fuzzy set theory are reviewed from [8], [17], [22], [36], and [38]. In Section III, we

Manuscript received November 19, 1997; revised July 15, 1998. This work was supported in part by the National Science Council, Republic of China, under Grant NSC 87-2213-E-009-011.

T.-P. Wu is with the Department of Computer and Information Science, National Chiao Tung University, Hsinchu, Taiwan, R.O.C.

S.-M. Chen is with the Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C.

Publisher Item Identifier S 1083-4419(99)00768-2.

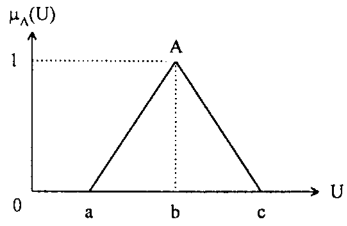

Fig. 1. The triangular fuzzy set A.

introduce how to use the α-cuts of the fuzzy equivalence relation to partition the numerical data and to derive membership functions of variables from them. And then we can directly induce fuzzy rules from the data partitions. In Section IV, we use an example to illustrate our new algorithm and the experimental results are shown. In Section V, we apply the proposed fuzzy learning algorithm to the Iris classification problem and compare the experimental results to that of Hong and Lee's algorithm [11]. The conclusions are discussed in Section VI.

II. Basic Concepts of Fuzzy Sets

The theory of fuzzy sets was proposed by Zadeh in 1965 [38]. A fuzzy set can be characterized by a membership function. Let U be the universe of discourse, $U = \{x_1, x_2, \cdots, x_n\}$. A fuzzy set A in the universe of discourse U is a set of ordered pairs $\{(x_1, \mu_A(x_1)), (x_2, \mu_A(x_2)), \cdots, (x_n, \mu(x_n))\}$, where μ_A is the membership function of the fuzzy set A, $\mu_A\colon U \rightarrow [0, 1]$, and $\mu_A(x_i)$ indicates the membership degree of x_i in the fuzzy set A. If the universe of discourse U is an infinite set, then the fuzzy set A can be represented by

$$A = \int_U \mu_A(x)/x, \qquad x \in U.$$

The triangular fuzzy set A shown in Fig. 1 can be represented by a triplet (a, b, c), where a and c are called the left vertex and right vertex of the triangular fuzzy set A, respectively.

The α-cut A_α of a fuzzy set A in the universe of discourse U is defined as follows:

$$A_\alpha = \{x | \mu_A(x) \geq \alpha, x \in U\}, \qquad \alpha \in [0, 1].$$

Based on the definition of α-cuts, the fuzzy set A can be represented as follows:

$$A = \bigcup_{\alpha \in [0, 1]} \alpha A_\alpha.$$

A fuzzy relation among fuzzy sets $A_1, A_2, \cdots, A_n$ is a subset of the Cartesian product $A_1 \times A_2 \times \cdots \times A_n$ and is denoted by $R(A_1, A_2, \cdots, A_n)$, where $A_1 \times A_2 \times \cdots \times A_n = \{(x_1, x_2, \cdots, x_n) | x_i \in A_i \text{ and } 1 \leq i \leq n\}$. The membership function of the fuzzy relation $R(A_1, A_2, \cdots, A_n)$ is represented by $\mu_R(x_1, x_2, \cdots, x_n)$, where $x_i \in A_i$ and $1 \leq i \leq n$. Let $R_1(A, B)$ and $R_2(B, C)$ be two binary fuzzy relations with a common fuzzy set B. The composition of $R_1(A, B)$ and $R_2(B, C)$ is denoted by $R_1(A, B) \circ R_2(B, C)$ and is defined as follows:

$$\mu_{R_1 \circ R_2}(x, z) = \max_{y \in B} \min[\mu_{R_1}(x, y), \mu_{R_2}(y, z)]$$

where for all $(x, y) \in A \times B$, $(y, z) \in B \times C$, $(x, z) \in A \times C$.

Let A be a fuzzy set of the universe of discourse U. A binary fuzzy relation $R(A, A)$ is reflexive if and only if $\mu_R(x, x) = 1, \forall x \in A$. A fuzzy relation $R(A, A)$ is symmetric if and only if $\mu_R(x, y) = \mu_R(y, x), \forall x, y \in A$. A fuzzy relation $R(A, A)$ is transitive if and only if $\mu_R(x, z) \geq \max_{y \in X} \min[\mu_R(x, y), \mu_R(y, z)], \forall, (x, z) \in A \times A$. The transitive closure of a binary fuzzy relation $R(A, A)$ is denoted by $R^T(A, A)$ and defined as follows:

$$R^T(A, A) = R^i$$

where $R^i = R^{i-1} \circ R$, $R^{i+1} = R^i$, and $i \geq 2$.

If a binary fuzzy relation R is reflexive and symmetric, then the binary fuzzy relation R is a compatibility relation. If a binary fuzzy relation R is reflexive, symmetric, and transitive, then the binary fuzzy relation R is called a fuzzy equivalence relation. Assume that A is a fuzzy set, then, the α-cuts of a binary fuzzy relation $R(A, A)$ can be defined as follows:

$$R_\alpha = \{(x, y) | \mu_R(x, y) \geq \alpha, (x, y) \in A \times A\}$$

where $\alpha \in [0, 1]$. Thus, every binary fuzzy relation R can be represented in terms of its α-cut R_α i.e.,

$$R = \bigcup_{\alpha \in [0, 1]} \alpha R_\alpha.$$

Furthermore, it can be easily shown that a fuzzy equivalence relation $R(A, A)$ can be considered to effectively group elements into equivalence classes whose members are similar to each other in some specified degree by taking an α-cut R_α, where $\alpha \in [0, 1]$. Each of these equivalence classes forms a partition of A. Two elements x, y of A belong to the same block of partition if and only if $R(x, y) \geq \alpha$.

III. A New Fuzzy Learning Algorithm

In an n-input-single-output fuzzy system, the fuzzy rules have the following general format:

$$R_j\colon \textbf{IF } X_1 \text{ is } A_{1,j} \textbf{ AND } X_2 \text{ is } A_{2,j} \textbf{ AND } \cdots \textbf{ AND } X_n \text{ is } A_{n,j} \textbf{ THEN } Y \text{ is } B_j$$

where the variables X_i $(i = 1, \cdots, n)$ appearing in the antecedent parts of the fuzzy rules R_j are called the input linguistic variables, the variable Y in the consequent part of the fuzzy rule R_j is called the output linguistic variable, the fuzzy sets $A_{i,j}$ are called the input fuzzy sets of the input linguistic variable X_i of the fuzzy rule R_j, and the fuzzy set B_j is called the output fuzzy set of the output linguistic variable Y of the fuzzy rule R_j.

In many real-world applications, the numerical data are easily obtained from the instruments or the environment. Thus, the fuzzy rules can be learned from these numerical data. Assume that there are m input–output pairs given as the numerical training data set P of the n-input-single-output fuzzy system

$$P = \{(x_{1,j}, \cdots, x_{n,j}, y_j) | j = 1, \cdots, m\} \quad (1)$$

where $x_{i,j}$ is the value of the ith input linguistic variable X_i of the jth input–output pair $(x_{1,j}, \cdots, x_{n,j}, y_j)$, and y_j is the value of the output linguistic variable Y of the jth input–output pair $(x_{1,j}, \cdots, x_{n,j}, y_j)$. The value $x_{i,j}$ of the input linguistic variable X_i is called the input value, and the value y_j of the output linguistic variable Y is called the output value, where $1 \leq i \leq n$ and $1 \leq j \leq m$. The input-value set of the input linguistic variable X_i consists of some input values of the input linguistic variable X_i in the input–output pairs of the training data set P, where $1 \leq i \leq n$. The output-value set of the output linguistic variable Y consists of some output values of the output linguistic variable Y in the input–output pairs of the training data set P.

For simplicity, we only consider to construct the multiple-inputs-single-output (MISO) fuzzy model. Since the outputs of the MISO fuzzy model are independent, the general rule structure of the MISO fuzzy model can easily be represented as a collection of the rules of the multiple-inputs-multiple-outputs (MIMO) fuzzy models [22].

As discussed in the previous section, a fuzzy equivalence relation is reflexive, symmetric, and transitive, and it can divide the crisp data into different groups by its α-cuts. Instead of finding the fuzzy equivalence relation directly, we can determine a fuzzy compatibility relation (reflexive and symmetric) in terms of an appropriate distance function applied to the given data. Then, a fuzzy equivalence relation can be obtained by the max–min transitive closure of the fuzzy compatibility relation. Before constructing the fuzzy equivalence relation of the training data set P, we first use the output values of the output linguistic variable Y as the key to sort the training data set P in an ascending order [11]. Then, the sorted training data set P' can be obtained as follows:

$$\begin{aligned} P' = \{&(x'_{1,p}, x'_{2,p}, \cdots, x'_{n,p}, y'_p) | y'_{p_1} \leq y'_{p_2}, \\ &p = 1, 2, \cdots, m, \quad \text{and} \quad 1 \leq p_1 \leq p_2 \leq m\} \end{aligned} \quad (2)$$

where $(x'_{1,p}, x'_{2,p}, \cdots, x'_{n,p}, y'_p) \in P$. We only consider the output values $y'_1, y'_2, \cdots, y'_m$ of the output linguistic variable Y in the sorted training data set P'. The fuzzy compatibility relation $R(y'_{p_1}, y'_{p_2})$ between the output values y'_{p_1} and y'_{p_2} of the output linguistic variable Y in the sorted training data set P' can be defined in terms of the Euclidean distance [17] as follows:

$$R(y'_{p_1}, y'_{p_2}) = 1 - \frac{|y'_{p_1} - y'_{p_2}|}{\delta} \quad (3)$$

where y'_{p_1} and y'_{p_2} are the output values of the output linguistic variable Y in the sorted training data set P', where $p_1 \in [1, m]$ and $p_2 \in [1, m]$, and δ is a constant which ensures the compatibility relation $R(y'_{p_1}, y'_{p_2}) \in [0, 1]$. That is, if $R(y'_{p_1}, y'_{p_2}) < 0$, then we let $R(y'_{p_1}, y'_{p_2}) = 0$. In this paper, we calculate the value of δ as follows:

$$\delta = \frac{\sum_{i=1}^{m-1} |y_i - y_m|}{m-1} \quad (4)$$

where y_m is the maximum value of the output linguistic variable Y in the sorted training data set P'. In general, the relation defined in (3) is a fuzzy compatibility relation, but it is not necessarily a fuzzy equivalence relation. The fuzzy equivalence relation $R^T(y'_{p_1}, y'_{p_2})$ between the output values y'_{p_1} and y'_{p_2} of the output linguistic variable Y in the sorted training data set P' can be obtained from the max–min transitive closure of the compatibility relation $R(y'_{p_1}, y'_{p_2})$ between the output values y'_{p_1} and y'_{p_2} of the output linguistic variable Y in the sorted training data set P'.

After the fuzzy equivalence relation $R^T(y'_{p_1}, y'_{p_2})$ between the output values of the output linguistic variable Y in the sorted training data set P' has been defined, we can divide the sorted training data set P' into different partitions based on the α-cuts of the fuzzy equivalence relation $R^T(y'_{p_1}, y'_{p_2})$. Assume that we partition the sorted training data set P' into r different subsets G_j $(j = 1, \cdots, r)$, and the jth subset G_j of the sorted training data P' can be represented as follows:

$$\begin{aligned} G_j = \{&(x'_{1,p}, x'_{2,p}, \cdots, x'_{n,p}, y'_p) | R^T(y'_{p_1}, y'_{p_2}) \geq \alpha, \\ &\alpha \in [0, 1], 1 \leq p \leq m, 1 \leq p_1 \leq m, \\ &\text{and} \quad 1 \leq p_2 \leq m\} \end{aligned} \quad (5)$$

where α is the threshold value that is chosen to divide the sorted training data set P' adequately, $1 \leq j \leq r$ and r is the number of subsets obtained from the sorted training data set P'.

Assume that the jth output-value set O_j of the output linguistic variable Y and the jth input-value set $I_{i,j}$ of the input linguistic variable X_i are obtained from the jth subset G_j of the sorted training data set P', then

$$\begin{aligned} O_j = \{&y_p | \forall (x_{1,p}, x_{2,p}, \cdots, x_{n,p}, y_p) \in G_j, 1 \leq p \leq m\}, \\ &1 \leq j \leq r \end{aligned} \quad (6)$$

$$\begin{aligned} I_{i,j} = \{&x_{i,p} | \forall (x_{1,p}, x_{2,p}, \cdots, x_{i,p}, \cdots, x_{n,p}, y_p) \in G_j, \\ &1 \leq p \leq m\}, \quad 1 \leq i \leq n, 1 \leq j \leq r. \end{aligned} \quad (7)$$

Thus, we can construct the membership functions of the output fuzzy sets of the output linguistic variable Y from the output values of the output-value set O_j, where $j = 1, \cdots, r$. Since the output values in the sorted training set P' have been divided into r different output-value sets O_j $(j = 1, \cdots, r)$ based on the α-cuts of the equivalence relation, each output-value set O_j of the output linguistic variable Y can be thought of as the α-cut $A_{j,\alpha}$ of the output fuzzy set A_j of the output linguistic variable Y. That is

$$A_{j,\alpha} = \{y | y \in O_j \quad \text{and} \quad \mu_{A_j}(y) \geq \alpha\}, \quad j = 1, \cdots, r \quad (8)$$

where $\mu_{A_j}(y)$ is the membership function of the output fuzzy set A_j of the output linguistic variable Y, and O_j is the jth output-value set of the output linguistic variable Y.

Because the output values of the α-cut $A_{j,\alpha}$ of the output linguistic variable Y are permuted in ascending order (i.e., $y_i \leq y_j$ if $i < j$), we can simply calculate the average of the minimum and the maximum output values of the α-cut set $A_{j,\alpha}$ and define it as the center of the output fuzzy set A_j of the output linguistic variable Y. That is

$$b_j = \frac{y_{\min} + y_{\max}}{2} \tag{9}$$

where b_j is the center of the output fuzzy set A_j; $y_{\min}$ and $y_{\max}$ are the minimum and the maximum elements of the α-cut $A_{j,\alpha}$, respectively. The membership grades of the minimum and maximum elements of the α-cut $A_{j,\alpha}$ are set to the threshold value α, where $\alpha \in [0, 1]$. According to these three points $(b_j, 1)$, $(y_{\min}, \alpha)$, and $(y_{\max}, \alpha)$, we can calculate the left vertex $(a_j, 0)$ and the right vertex $(c_j, 0)$ of the output fuzzy set A_j of the output linguistic variable Y by interpolation as follows:

$$a_j = b_j - \frac{b_j - y_{\min}}{1-\alpha} \tag{10}$$

$$c_j = b_j + \frac{y_{\max} - b_j}{1-\alpha} \tag{11}$$

where b_j is obtained using (9), $y_{\min}$ and $y_{\max}$ are the minimum and the maximum elements of the α-cut $A_{j,\alpha}$, respectively. Thus, the membership function $\mu_{A_j}(y)$ of the output fuzzy set A_j of the output linguistic variable Y can be represented by the triplet (a_j, b_j, c_j) as follows:

$$\mu_{A_j}(y) = \begin{cases} \dfrac{y - a_j}{|b_j - a_j|}, & \text{if } a_j \leq y \leq b_j, \\ \dfrac{c_j - y}{|c_j - b_j|}, & \text{if } b_j \leq y \leq c_j, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

For the input linguistic variable X_i, its corresponding input values in the sorted training data set P' also have concurrently been divided into r input-value sets $I_{i,1}, I_{i,2}, \cdots, I_{i,r}$ based on the α-cuts of the equivalence relation between the output values of the output linguistic variable Y in the sorted training data set P'. In the jth input-value set $I_{i,j}$ of the input linguistic variable X_i, we also sort the input values in the ascending order. The fuzzy equivalence relation between the input values of the jth input-value set $I_{i,j}$ of the input linguistic variable X_i can be constructed based on the similarity between the input values. We also define the fuzzy compatibility relation between the input values of the input-value set $I_{i,j}$ of the input linguistic variable X_i in terms of the Euclidean distance [17]

$$R_{i,j}(x_{i,p_1}, x_{i,p_2}) = 1 - \frac{|x_{i,p_1} - x_{i,p_2}|}{\delta} \tag{13}$$

where $x_{i,p_1} \in I_{i,j}$, $x_{i,p_2} \in I_{i,j}$, $1 \leq i \leq n$, $1 \leq j \leq r$, $1 \leq p_1 \leq |I_{i,j}|$, and $1 \leq p_2 \leq |I_{i,j}|$. $|I_{i,j}|$ denotes the number of the elements of the input-value set $I_{i,j}$. The value of δ is calculated as follows:

$$\delta = \frac{\sum_{p=1}^{|I_{i,j}|-1} |x_{i,p} - x_{i,\max}|}{|I_{i,j}| - 1} \tag{14}$$

where $|I_{i,j}|$ denotes the number of the elements of the input-value set $I_{i,j}$ of the input linguistic variable X_i, $x_{i,\max}$ is the maximum value of the input linguistic variable X_i in the input-value set $I_{i,j}$, and $x_{i,p} \in I_{i,j}$. By the max–min transitive closure of the compatibility relation $R_{i,j}(x_{i,p_1}, x_{i,p_2})$, we can obtain the fuzzy equivalence relation $R^T_{i,j}(x_{i,p_1}, x_{i,p_2})$ between the input values of the input-value set $I_{i,j}$ of the input linguistic variable X_i. Furthermore, based on the α-cuts of the fuzzy equivalence relation, the jth input-value set $I_{i,j}$ of the input linguistic variable X_i can be divided into different "input-value subsets" $I_{i,j,k}$ of the jth input-value set $I_{i,j}$ of the input linguistic variable X_i as follows:

$$\begin{aligned} I_{i,j,k} = \{x_{i,p} | x_{i,p} \in I_{i,j},\ & R^T_{i,j}(x_{i,p_1}, x_{i,p_2}) \geq \alpha, \\ & \alpha \in [0, 1],\ 1 \leq p \leq |I_{i,j}|,\ 1 \leq p_1 \leq |I_{i,j}|, \\ & \text{and} \quad 1 \leq p_2 \leq |I_{i,j}|\},\ k = 1, \cdots, T_j(X_i) \end{aligned} \tag{15}$$

where $T_j(X_i)$ is the number of the input-value subsets obtained from the jth input-value set $I_{i,j}$ of the input linguistic variable X_i based on the α-cuts of the fuzzy equivalence relation between the input values of the jth input-value set $I_{i,j}$ of the input linguistic variable X_i. $|I_{i,j}|$ denotes the number of the elements of the input-value set $I_{i,j}$.

Every input-value subset $I_{i,j,k}$ of the input linguistic variable X_i can also be thought of as the α-cut $A_{(i,j,k),\alpha}$ of the corresponding input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i. Then, the α-cut $A_{(i,j,k),\alpha}$ of the input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i is defined as follows:

$$\begin{aligned} A_{(i,j,k),\alpha} = \{x_{i,p} | x_{i,p} \in I_{i,j,k},\ \mu_{A_{i,j,k}}(x_{i,p}) \geq \alpha\}, \\ 1 \leq j \leq r,\ 1 \leq k \leq T_j(X_i) \end{aligned} \tag{16}$$

where $\mu_{A_{i,j,k}}(x_i)$ is the membership function of the input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i, α is the threshold value, and $\alpha \in [0, 1]$. The input values in the α-cut $A_{(i,j,k),\alpha}$ are also permuted in ascending order (i.e., $x_{i,p_1} \leq x_{i,p_2}$, if $p_1 \leq p_2$). The membership function $\mu_{A_{i,j,k}}(x_i)$ of the input fuzzy set $A_{i,j,k}$ can be represented by a triplet $(a_{i,j,k}, b_{i,j,k}, c_{i,j,k})$, where $b_{i,j,k}$, $a_{i,j,k}$, and $c_{i,j,k}$ are the center, the left vertex, and the right vertex of the input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i, respectively. Then, the triplet $(a_{i,j,k}, b_{i,j,k}, c_{i,j,k})$ can be calculated by the minimum and the maximum elements of the α-cut $A_{(i,j,k),\alpha}$ of the input fuzzy set $A_{i,j,k}$ and by interpolation as follows:

$$b_{i,j,k} = \frac{x_{i,\min} + x_{i,\max}}{2} \tag{17}$$

$$a_{i,j,k} = b_{i,j,k} - \frac{b_{i,j,k} - x_{i,\min}}{1-\alpha} \tag{18}$$

$$c_{i,j,k} = b_{i,j,k} + \frac{x_{i,\max} - b_{i,j,k}}{1-\alpha} \tag{19}$$

where $x_{i,\min}$ and $x_{i,\max}$ are the minimum and the maximum elements of the α-cut $A_{(i,j,k),\alpha}$ of the input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i, $1 \leq j \leq r$, $1 \leq k \leq T_j(X_i)$, and $T_j(X_i)$ is the number of the input-value subsets $I_{i,j,k}$ obtained from the jth input-value set $I_{i,j}$ of the input linguistic variable X_i. Furthermore, the membership function $\mu_{A_{i,j,k}}(x_i)$ of the input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i can be represented by the triplet $(a_{i,j,k}, b_{i,j,k}, c_{i,j,k})$ as follows:

$$\mu_{A_{i,j,k}}(x_i) = \begin{cases} \dfrac{x_i - a_{i,j,k}}{|b_{i,j,k} - a_{i,j,k}|}, & \text{if } a_{i,j,k} \le x_i \le b_{i,j,k}, \\ \dfrac{c_{i,j,k} - x_i}{|c_{i,j,k} - b_{i,j,k}|}, & \text{if } b_{i,j,k} \le x_i \le c_{i,j,k}, \\ 0, & \text{otherwise,} \end{cases} \tag{20}$$

where $b_{i,j,k}$ and $a_{i,j,k}$, and $c_{i,j,k}$ are the center, the left vertex, and the right vertex of the input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i, respectively, $1 \le j \le r$, $1 \le k \le T_j(X_i)$, and $T_j(X_i)$ is the number of the input-value subset $I_{i,j,k}$ obtained from the jth input-value set $I_{i,j}$ of the input linguistic variable X_i.

Based on (3)–(20), we can partition the input–output pairs of the sorted training data set P' into the different input-value subset $I_{i,j,k}$ of the input linguistic variable X_i and the output-value set O_j of the output linguistic variable Y, where $1 \le i \le n$, $1 \le j \le r$, and $1 \le k \le T_j(X_i)$. We also can derive the membership function $\mu_{A_{i,j,k}}(x_i)$ of the input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i and the membership function $\mu_{A_j}(y)$ of the output fuzzy set A_j of the output linguistic variable Y from the data residing in the input-value subset $I_{i,j,k}$ of the input linguistic variable X_i and the output-value set O_j of the output linguistic variable Y, respectively.

After deriving the membership functions of the input linguistic variables and the output linguistic variables, we continue to generate the fuzzy rules based on the hierarchical relationships between the input-value subsets of the input linguistic variables and the output-value sets of the output linguistic variables. In the following, we take a system with two input linguistic variables X_1, X_2, and one output linguistic variable Y as an example to illustrate the hierarchical relationships between the output-value sets and the input-value subsets.

First, the jth output-value set O_j of the output linguistic variable Y and the jth input-value set $I_{i,j}$ of the input linguistic variable X_i are obtained based on the α-cuts of the equivalence relation between the output values of the output linguistic variable Y in the training data set P', where $i = 1, 2$; $1 \le j \le r$, and r is the number of subsets obtained from the sorted training data set P'. Furthermore, the kth input-value subset $I_{i,j,k}$ of the input linguistic variable X_i can be obtained based on the α-cuts of the equivalence relation between the input values of the jth input-value set $I_{i,j}$ of the input linguistic variable X_i, where $i = 1, 2$; $1 \le j \le r$, $1 \le k \le T_j(X_i)$, and $T_j(X_i)$ is the number of input-value subsets $I_{i,j,k}$ obtained from the jth input-value set $I_{i,j}$ of the input linguistic variable X_i.

There exists the hierarchical relationship between the jth output-value set O_j of the output linguistic variable Y and the kth input-value subsets $I_{1,j,k}$ of the input linguistic variable X_1, where $1 \le j \le r$ and $1 \le k \le T_j(X_1)$. Assume that the input–output pair $(x_{1,p}, x_{2,p}, y_p)$ belongs to the sorted training data set P', where $1 \le p \le m$. If the input value $x_{1,p}$ of the input linguistic variable X_1 belongs to the input-value subset $I_{i,j,k}$, then we can infer that the corresponding output value y_p of the output linguistic variable Y should belong to the output-value set O_j. In Fig. 2, we can see the hierarchical relationship between the jth output-value sets O_j of the output linguistic variable Y and the kth input-value subset $I_{1,j,k}$ of the input linguistic variable X_1. Hence, we can obtain the hierarchical relationship between the jth output fuzzy set A_j of the output linguistic variable Y and the kth input fuzzy set $A_{1,j,k}$ of the input linguistic variable X_1 based on the hierarchical relationship between the jth output-value set O_j of the output linguistic variable Y and the kth input-value subset $I_{1,j,k}$ of the input linguistic variable X_1. The hierarchical relationship between the jth output fuzzy set A_j of the output linguistic variable Y and the kth input fuzzy set $A_{1,j,k}$ of the input linguistic variable X_1 is shown in Fig. 3, where $1 \le j \le r$ and $1 \le k \le T_j(X_1)$.

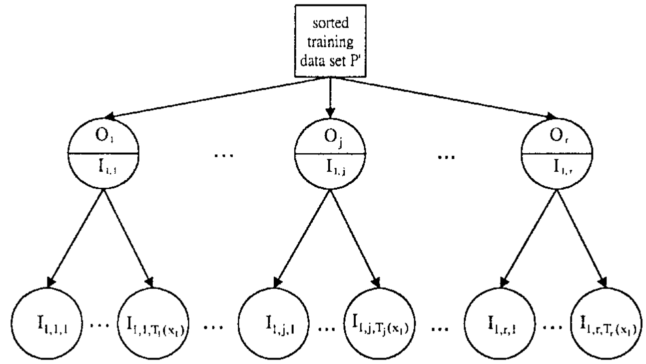

Fig. 2. The relationship between the output-value set O_j of the output linguistic variable Y and the input-value subset $I_{1,j,k}$ of the input linguistic variable X_1.

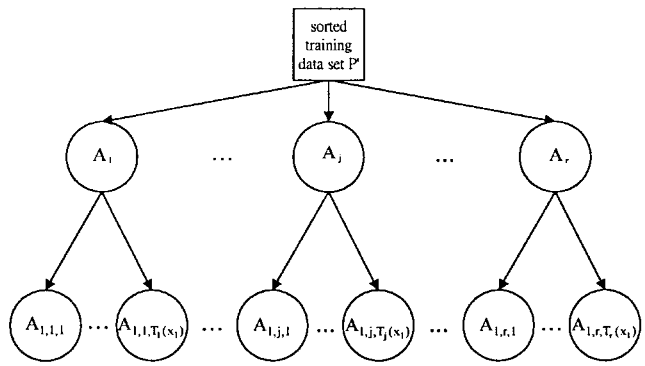

Fig. 3. The relationship between the output fuzzy set A_j of the output linguistic variable Y and the input fuzzy set $A_{1,j,k}$ of the input linguistic variable X_1.

From Fig. 3, the fuzzy rules can be generated based on the hierarchical relationship between the output fuzzy set A_j of the output linguistic variable Y and the input fuzzy set $A_{1,j,k}$

of the input linguistic variable X_1 as follows:

$$\begin{array}{ll}\textbf{IF}\, X_1 \text{ is } A_{1,1,1} & \textbf{THEN}\, Y \text{ is } A_1\\ & \vdots \\ \textbf{IF}\, X_1 \text{ is } A_{1,1,T_1(x_1)} & \textbf{THEN}\, Y \text{ is } A_1\\ & \vdots \\ \textbf{IF}\, X_1 \text{ is } A_{1,j,1} & \textbf{THEN}\, Y \text{ is } A_j\\ & \vdots \\ \textbf{IF}\, X_1 \text{ is } A_{1,j,T_j(x_1)} & \textbf{THEN}\, Y \text{ is } A_j\\ & \vdots \\ \textbf{IF}\, X_1 \text{ is } A_{1,r,1} & \textbf{THEN}\, Y \text{ is } A_r\\ & \vdots \\ \textbf{IF}\, X_1 \text{ is } A_{1,r,T_r(x_1)} & \textbf{THEN}\, Y \text{ is } A_r.\end{array} \tag{21}$$

By the same way, we can also obtain the hierarchical relationship between the jth output-value set O_j of the output linguistic variable Y and the kth input-value subset $I_{2,j,k}$ of the input linguistic variable X_2 as shown in Fig. 4, where $1 \le j \le r$ and $1 \le k \le T_j(X_2)$. The hierarchical relationship between the jth output fuzzy set A_j of the output linguistic variable Y and the kth input fuzzy set $A_{2,j,k}$ of the input linguistic variable X_2 is shown in Fig. 5, where $1 \le j \le r$ and $1 \le k \le T_j(X_2)$. From Fig. 5, the fuzzy rules can be generated based on the hierarchical relationship between the output fuzzy set A_j of the output linguistic variable Y and the input fuzzy sets $A_{2,j,k}$ of the input linguistic variable X_2 as follows:

$$\begin{array}{ll}\textbf{IF}\, X_2 \text{ is } A_{2,1,1} & \textbf{THEN}\, Y \text{ is } A_1\\ & \vdots \\ \textbf{IF}\, X_2 \text{ is } A_{2,1,T_1(x_2)} & \textbf{THEN}\, Y \text{ is } A_1\\ & \vdots \\ \textbf{IF}\, X_2 \text{ is } A_{2,j,1} & \textbf{THEN}\, Y \text{ is } A_j\\ & \vdots \\ \textbf{IF}\, X_2 \text{ is } A_{2,j,T_j(x_2)} & \textbf{THEN}\, Y \text{ is } A_j\\ & \vdots \\ \textbf{IF}\, X_2 \text{ is } A_{2,r,1} & \textbf{THEN}\, Y \text{ is } A_r\\ & \vdots \\ \textbf{IF}\, X_2 \text{ is } A_{2,r,T_r(x_2)} & \textbf{THEN}\, Y \text{ is } A_r.\end{array} \tag{22}$$

In fact, we can see that the fuzzy rules whose antecedent part is "X_1 is $A_{1,j,k}$" in (21) and the fuzzy rules whose antecedent part is "X_2 is $A_{2,j,k}$" in (22) share the same consequent part "Y is A_j." For a training input–output pair $(x_{1,p}, x_{2,p}, y_p)$ of the sorted training data set P', if the input value $x_{1,p}$ belongs to the input-value set $I_{1,j}$ of the input linguistic variable X_1 and if the input value $x_{2,p}$ belongs to the input-value set $I_{2,j}$ of the input linguistic variable X_2, then the output value y_p belongs to the output-value set O_j of

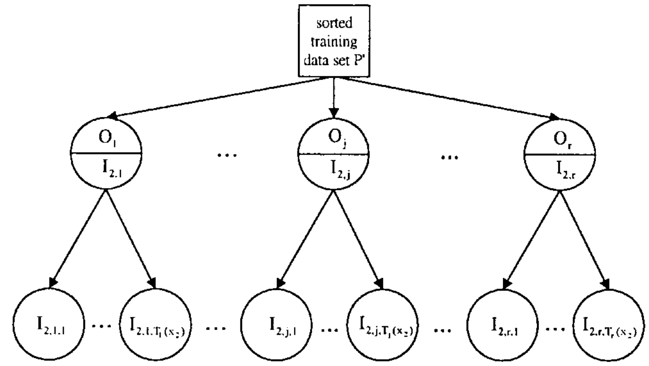

Fig. 4. The relationship between the output-value set O_j of the output linguistic variable Y and the input-value subset $I_{2,j,k}$ of the input linguistic variable X_2.

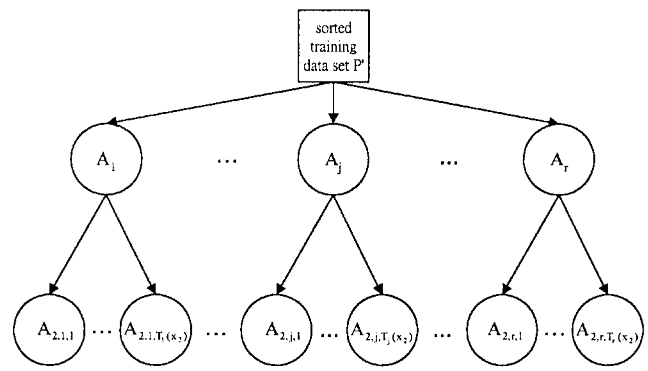

Fig. 5. The relationship between the output fuzzy set A_j of the output linguistic variable Y and the input fuzzy set $A_{2,j,k}$ of the input linguistic variable X_2.

the output linguistic variable Y. Thus, if the input–output pair $(x_{1,p}, x_{2,p}, y_p)$ satisfies both the antecedent parts of "X_1 is A_{1,j,k_1}" and "X_2 is A_{2,j,k_2}" for some $k_1 \in [1, T_j(X_1)]$ and some $k_2 \in [1, T_j(X_2)]$, then we can infer that the consequent part of the fuzzy rule is "Y is A_j," where $p \in [1, m]$, $T_j(X_i)$ is the number of the input-value subsets obtained from the jth input-value set $I_{i,j}$ of the input linguistic variable X_i, where $i = 1, 2$.

The hierarchical relationship between the output fuzzy set A_j of the output linguistic variable Y, the input fuzzy set A_{1,j,k_1} of the input linguistic variable X_1, and the input fuzzy set A_{2,j,k_2} of the input linguistic variable X_2 can be obtained as shown in Fig. 6, where $k_1 \in [1, T_j(X_1)]$ and $k_2 \in [1, T_j(X_2)]$.

The fuzzy rules in (21) and (22) can be combined based on the hierarchical relationship between the output fuzzy set A_j of output linguistic variable Y, the input fuzzy set A_{1,j,k_1} of the input linguistic variable X_1, and the input fuzzy set A_{2,j,k_2} of the input linguistic variable X_2 as follows:

$$\textbf{IF}\, X_1 \text{ is } A_{1,j,k_1} \,\textbf{AND}\, X_2 \text{ is } A_{2,j,k_2} \,\textbf{THEN}\, Y \text{ is } A_j \tag{23}$$

where $1 \le j \le r$, $1 \le k_1 \le T_j(X_1)$, and $1 \le k_2 \le T_j(X_2)$.

Finally, we can generate $\sum_{j=1}^{r} T_j(x_1)T_j(x_2)$ fuzzy rules from the given numerical training data set. In general, if

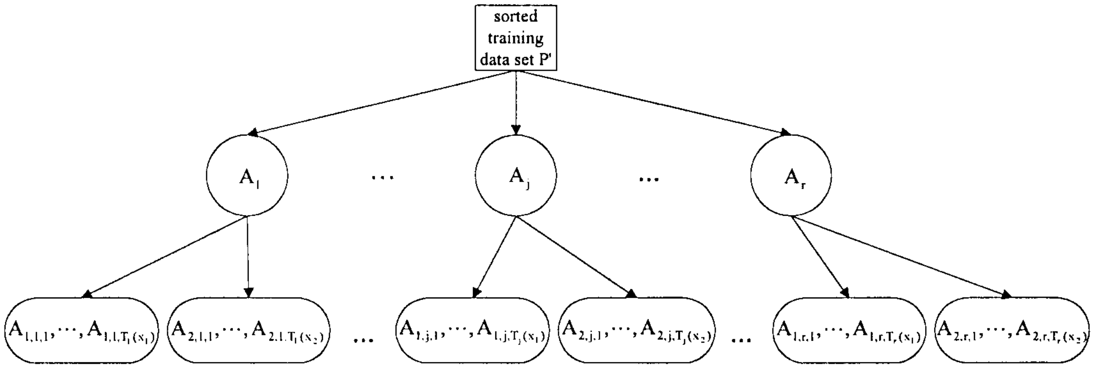

Fig. 6. Relationship between the output fuzzy set A_j of the output linguistic variable Y, the input fuzzy set A_{1,j,k_1} of the input linguistic variable X_1 and the input fuzzy set A_{2,j,k_2} of the input linguistic variable X_2.

there are n input linguistic variables $X_1, \cdots, X_n$ and a single output linguistic variable Y in a fuzzy system, then the total number of fuzzy rules generated from the training data set P based on the proposed fuzzy learning algorithm is

$$\sum_{j=1}^{r} \prod_{i=1}^{n} T_j(X_i) \tag{24}$$

where r is the number of output fuzzy sets of the output linguistic variable Y, and $T_j(X_i)$ is the number of input fuzzy sets of the input linguistic variable X_i.

However, it is necessary to check whether some of the generated fuzzy rules are unnecessary or some input fuzzy sets of the input linguistic variables in the antecedent parts of the fuzzy rules are equivalent or redundant. If there are equivalent input fuzzy sets of the same input linguistic variable in the antecedent parts of the generated fuzzy rules, we need to perform the merge operation to simplify the generated fuzzy rules.

Consider the following two fuzzy rules with two input linguistic variables X_1, X_2, and one output linguistic variable Y:

$$\textbf{IF } X_1 \text{ is } A_{1,1,1} \textbf{ AND } X_2 \text{ is } A_{2,1,1} \textbf{ THEN } Y \text{ is } A_1$$
$$\textbf{IF } X_1 \text{ is } A_{1,2,1} \textbf{ AND } X_2 \text{ is } A_{2,2,1} \textbf{ THEN } Y \text{ is } A_2$$

where A_1 and A_2 are the output fuzzy sets of the output linguistic variable Y, $A_{1,1,1}$ and $A_{1,2,1}$ are the input fuzzy sets of the input linguistic variable X_1, and $A_{2,1,1}$ and $A_{2,2,1}$ are the input fuzzy sets of the input linguistic variable X_2. The input fuzzy sets $A_{1,1,1}$ and $A_{1,2,1}$ of the input linguistic variable X_1, and the input fuzzy sets $A_{2,1,1}$ and $A_{2,2,1}$ of the input linguistic variable X_2 are shown in Fig. 7.

From Fig. 7, we can see that the similarity degree between the input fuzzy sets $A_{2,1,1}$ and $A_{2,2,1}$ of the input linguistic variable X_2 is high. Since the two input fuzzy sets seem to be equivalent, we can generate a new membership function of the new input fuzzy set $A_{2,new}$ by merging the membership functions of the input fuzzy sets $A_{2,1,1}$ and $A_{2,2,1}$ of the input linguistic variable X_2. After the new input fuzzy set $A_{2,new}$ of the input linguistic variable X_2 has been generated, we can replace the input fuzzy sets $A_{2,1,1}$ and $A_{2,2,1}$ in the antecedent parts of the fuzzy rules by $A_{2,new}$. The simplified fuzzy rules are listed as follows:

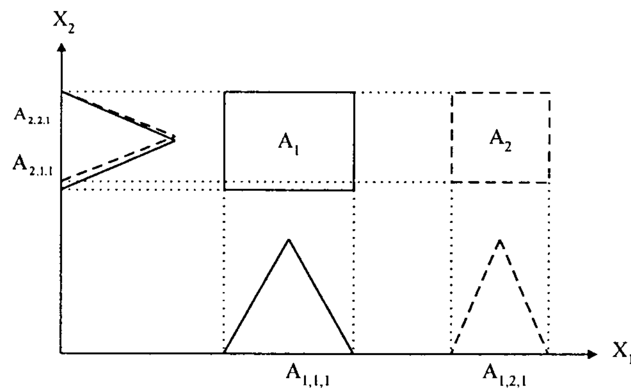

Fig. 7. The similarity between input fuzzy sets.

$$\textbf{IF } X_1 \text{ is } A_{1,1,1} \textbf{ AND } X_2 \text{ is } A_{2,new} \textbf{ THEN } Y \text{ is } A_1$$
$$\textbf{IF } X_1 \text{ is } A_{1,2,1} \textbf{ AND } X_2 \text{ is } A_{2,new} \textbf{ THEN } Y \text{ is } A_2.$$

There are a lot of methods to measure the similarity or equality degree between two distinct fuzzy sets, such as in [4], [5], [22], [31], and [37]. In this paper, we build the equality relation matrices for the input fuzzy sets of the input linguistic variable X_i based on the method proposed by Lin [22]. The equality degree of two distinct input fuzzy sets A_{i,j_1,k_1} and A_{i,j_2,k_2} of the input linguistic variable X_i is defined as follows:

$$E(A_{i,j_1,k_1}, A_{i,j_2,k_2}) = \frac{|A_{i,j_1,k_1} \cap A_{i,j_2,k_2}|}{|A_{i,j_1,k_1} \cup A_{i,j_2,k_2}|} \tag{25}$$

where $1 \le j_1 \le r$, $1 \le j_2 \le r$, $1 \le k_1 \le T_{j_1}(x_i)$, and $1 \le k_2 \le T_{j_2}(x_i)$. $|A_{i,j_1,k_1} \cap A_{i,j_2,k_2}|$ is the area of the intersection of the two distinct input fuzzy sets A_{i,j_1,k_1} and A_{i,j_2,k_2}. $|A_{i,j_1,k_1} \cup A_{i,j_2,k_2}|$ is the area of the union of the two distinct input fuzzy sets A_{i,j_1,k_1} and A_{i,j_2,k_2}.

By means of the α-cuts of the equality, we can obtain those fuzzy sets whose equality degree is higher than the specified threshold value α_{Equality}, where $\alpha_{\text{Equality}} \in [0, 1]$ and then merge their membership functions to generate a new one. Assume that $E(A_{i,j_1,k_1}, A_{i,j_2,k_2})$ is greater than the threshold value α_{Equality}, where the value of α_{Equality} is specified by the user and $\alpha_{\text{Equality}} \in [0, 1]$, then we can merge the two input fuzzy sets A_{i,j_1,k_1} and A_{i,j_2,k_2} of the input linguistic variable X_i to generate the new input fuzzy set $A_{i,new}$ which can be represented by the triplet $(a_{i,new}, b_{i,new}, c_{i,new})$. The triplet $(a_{i,new}, b_{i,new}, c_{i,new})$ is calculated by averaging the triplets $(a_{i,j_1,k_1}, b_{i,j_1,k_1}, c_{i,j_1,k_1})$ of the input fuzzy set A_{i,j_1,k_1} and the triplet $(a_{i,j_2,k_2}, b_{i,j_2,k_2}, c_{i,j_2,k_2})$ of the input fuzzy set A_{i,j_2,k_2}. Hence, the triplet $(a_{i,new}, b_{i,new}, c_{i,new})$ of the merged input fuzzy set $A_{i,new}$ can be obtained as follows:

$$a_{i,new} = \frac{a_{i,j_1,k_1} + a_{i,j_2,k_2}}{2} \tag{26}$$

$$b_{i,new} = \frac{b_{i,j_1,k_1} + b_{i,j_2,k_2}}{2} \tag{27}$$

$$c_{i,new} = \frac{c_{i,j_1,k_1} + c_{i,j_2,k_2}}{2}. \tag{28}$$

The membership function $\mu_{A_{i,new}}(x_i)$ of the input fuzzy set $A_{i,new}$ of the input linguistic variable X_i can be represented by the triplet $(a_{i,new}, b_{i,new}, c_{i,new})$ as follows:

$$\mu_{A_{i,new}}(x_i) = \begin{cases} \dfrac{x_i - a_{i,new}}{|b_{i,new} - a_{i,new}|}, & \text{if } a_{i,new} \le x_i \le b_{i,new}, \\ \dfrac{c_{i,new} - x_i}{|c_{i,new} - b_{i,new}|}, & \text{if } b_{i,new} \le x_i \le c_{i,new}, \\ 0, & \text{otherwise.} \end{cases} \tag{29}$$

The new fuzzy set of the input linguistic variable X_i can replace those input fuzzy sets of the input linguistic variable X_i whose equality degree is higher than the threshold value α_{Equality} in the antecedent parts of the fuzzy rules. Therefore, the number of the input fuzzy sets of the input linguistic variable X_i in the generated fuzzy rules is reduced, and we can obtain simpler and more efficient fuzzy rules.

Based on the method discussed previously, the proposed fuzzy learning algorithm which constructs the membership functions of the input linguistic variables and the output linguistic variable and generates fuzzy rules from the numerical training data set is stated as follows.

Input: The training data set P contains m input–output pairs $(x_{1,p}, \cdots, x_{n,p}, y_p)$ of the input linguistic variables $X_1, \cdots, X_n$, and the output linguistic variable Y, where $1 \le p \le m$.

Step 1: Sort the training data set P in ascending order based on the output values of the output linguistic variable Y and obtain the sorted training data set P'.

Step 2: Construct the equivalence relation R^T between the output values of the output linguistic variable Y in the sorted training data set P'.

Step 3: Divide the sorted training data in the sorted training set P' into r different output-value sets O_j of the output linguistic variable Y and r different input-value sets $I_{i,j}$ of the input linguistic variable X_i based on the α-cuts of the equivalence relations R^T derived from the output values of the output linguistic variable Y in the sorted training set P', where $1 \le i \le n$ and $1 \le j \le r$.

Step 4: Derive the membership function $\mu_{A_j}(y)$ of the output fuzzy set A_j of the output linguistic variable Y based on the α-cuts $A_{j,\alpha}$ of the output fuzzy set A_j with respect to the output-value sets O_j, where $1 \le j \le r$.

Step 5: Sort the input values of the input-value set $I_{i,j}$ of the input linguistic variable X_i in ascending order, where $1 \le i \le n$ and $1 \le j \le r$.

Step 6: Construct the equivalence relation $R^T_{i,j}$ between the input values of the jth input-value set $I_{i,j}$ of the input linguistic variable X_i, where $1 \le i \le n$ and $1 \le j \le r$.

Step 7: Divide the input values of the input-value set $I_{i,j}$ into $T_j(X_i)$ input-value subsets $I_{i,j,k}$ based on the α-cuts of the equivalence relation $R^T_{i,j}$, where $1 \le i \le n$, $1 \le j \le r$, $1 \le k \le T_j(X_i)$, and $T_j(X_i)$ is the number of input-value subsets obtained from the jth input-value set $I_{i,j}$ of the input linguistic variable X_i.

Step 8: Derive the input membership function $\mu_{A_{i,j,k}}(x_i)$ of the input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i based on the α-cut $A_{(i,j,k),\alpha}$ of the input fuzzy set $A_{i,j,k}$ with respect to the input-value subset $I_{i,j,k}$ of the input linguistic variable X_i, where $1 \le i \le n$ and $1 \le j \le r$, $1 \le k \le T_j(X_i)$, and $T_j(X_i)$ is the number of input-value subsets obtained from the jth input-value set $I_{i,j}$ of the input linguistic variable X_i.

Step 9: Generate the fuzzy rules based on the hierarchical relationship between the output fuzzy set A_j of the output linguistic variable Y and the corresponding input fuzzy set $A_{i,j,k}$ of the input linguistic variable X_i, where $1 \le i \le n$, $1 \le j \le r$, and $1 \le k \le T_j(X_i)$.

Step 10: Calculate the equality degree $E(A_{i,j_1,k_1}, A_{i,j_2,k_2})$ between the input fuzzy sets A_{i,j_1,k_1} and A_{i,j_2,k_2} of the input linguistic variable X_i based on (25), where $1 \le j_1 \le r$, $1 \le j_2 \le r$, $1 \le k_1 \le T_{j_1}(x_i)$, and $1 \le k_2 \le T_{j_2}(x_i)$.

Step 11: If $E(A_{i,j_1,k_1}, A_{i,j_2,k_2}) > \alpha_{\text{Equality}}$, where α_{Equality} is specified by the user, $\alpha_{\text{Equality}} \in [0, 1]$, $1 \le j_1 \le r$, $1 \le j_2 \le r$, $1 \le k_1 \le T_{j_1}(x_i)$, and $1 \le k_2 \le T_{j_2}(x_i)$ then
{
Construct the new input fuzzy set $A_{i,new}$ of the input linguistic variable X_i by merging the input fuzzy sets A_{i,j_1,k_1} and A_{i,j_2,k_2} of the input linguistic variable X_i;

Replace the input fuzzy sets A_{i,j_1,k_1} and A_{i,j_2,k_2} of the input linguistic variable X_i in the antecedent parts of the fuzzy rules by the new input fuzzy set $A_{i,new}$

}

IV. An Example

In the following, we use the example shown in [11] to demonstrate the proposed fuzzy learning algorithm for constructing membership functions and generating fuzzy rules from the numerical training data set.

Example 4.1: An insurance company evaluates the insurance fee of a person based on the age and the property of the person. Assume that the eight training examples are available as shown in Table I. Each input–output pair consists of two input linguistic variables, "Age" and "Property," and one output linguistic variable, "Insurance Fee." The unit of the property is ten thousand dollars. We will learn the membership functions of the two input linguistic variables "Age" and "Property," and the output linguistic variable "Insurance Fee" and generate fuzzy rules from the eight training samples based on the proposed algorithm.

TABLE I
The Training Data Set of the Insurance Fee Problem [11]

Age	Property	Insurance Fee
20	30	2000
25	30	2100
30	10	2200
45	50	2500
50	30	2600
60	10	2700
80	30	3200
80	40	3300

[Step 1]: /* Sort the training data set in ascending order according to the output values of the output linguistic variable */

Since the output values of the output linguistic variable "Insurance Fee" of the training data set shown in Table I have been arranged in ascending order, by performing Step 1 of the algorithm, the sorted training data set P' is shown as follows:

$$P' = \{(20, 30, 2000), (25, 30, 2100), (30, 10, 2200), (45, 50, 2500), (50, 30, 2600), (60, 10, 2700), (80, 30, 3200), (80, 40, 3300)\}.$$

[Step 2]: /* Construct the equivalence relation R^T between the output values of the output linguistic variable */

According to (3), the compatibility relation R between the output values of the output linguistic variable "Insurance Fee" is constructed as shown in (30) and (31), at the bottom of the page, where the equivalence relation R^T between the output values of the output linguistic variable "Insurance Fee" in the sorted training data set P' can be derived by the max–min transitive closure of the compatibility relation R as shown in (32), at the bottom of the page.

[Step 3]: /* Divide input–output pairs of the sorted training data set P' into different subsets based on the

$$R = \begin{bmatrix} 1.0 & 0.88 & 0.76 & 0.4 & 0.28 & 0.16 & 0 & 0 \\ 0.88 & 1.0 & 0.88 & 0.52 & 0.4 & 0.28 & 0 & 0 \\ 0.76 & 0.88 & 1.0 & 0.64 & 0.52 & 0.4 & 0 & 0 \\ 0.4 & 0.52 & 0.64 & 1.0 & 0.88 & 0.76 & 0.16 & 0.03 \\ 0.28 & 0.4 & 0.52 & 0.88 & 1.0 & 0.88 & 0.28 & 0.16 \\ 0.16 & 0.28 & 0.4 & 0.76 & 0.88 & 1.0 & 0.4 & 0.28 \\ 0 & 0 & 0 & 0.16 & 0.28 & 0.4 & 1.0 & 0.88 \\ 0 & 0 & 0 & 0.03 & 0.16 & 0.28 & 0.88 & 1.0 \end{bmatrix} \tag{30}$$

$$\delta = \frac{(3300-2000)+(3300-2100)+(3300-2200)+(3300-2500)+(3300-2600)+(3300-2700)+(3300-3200)}{7} = 828.57 \tag{31}$$

$$R^T = \begin{bmatrix} 1.0 & 0.88 & 0.88 & 0.64 & 0.64 & 0.64 & 0.4 & 0.4 \\ 0.88 & 1.0 & 0.88 & 0.64 & 0.64 & 0.64 & 0.4 & 0.4 \\ 0.88 & 0.88 & 1.0 & 0.64 & 0.64 & 0.64 & 0.4 & 0.4 \\ 0.64 & 0.64 & 0.64 & 1.0 & 0.88 & 0.88 & 0.4 & 0.4 \\ 0.64 & 0.64 & 0.64 & 0.88 & 1.0 & 0.88 & 0.4 & 0.4 \\ 0.64 & 0.64 & 0.64 & 0.88 & 0.88 & 1.0 & 0.4 & 0.4 \\ 0.4 & 0.4 & 0.4 & 0.4 & 0.4 & 0.4 & 1.0 & 0.88 \\ 0.4 & 0.4 & 0.4 & 0.4 & 0.4 & 0.4 & 0.88 & 1.0 \end{bmatrix}. \tag{32}$$

α-cuts of the equivalence relation between the output values of the output linguistic variable */

Assume that the threshold value α is 0.8, then the sorted training set P' can be divided into three subsets G_1, G_2, and G_3 based on the 0.8-cut of the equivalence relation R^T as follows:

$$\begin{aligned} G_1 &= \{(20, 30, 2000), (25, 30, 2100), \\ &\quad (30, 10, 2200)\} \\ G_2 &= \{(45, 50, 2500), (50, 30, 2600), \\ &\quad (60, 10, 2700)\} \\ G_3 &= \{(80, 30, 3200), (80, 40, 3300)\}. \end{aligned}$$

Furthermore, the output-value sets O_1, O_2, and O_3 of the output linguistic variable "Insurance Fee," the input-value sets $I_{1,1}$, $I_{1,2}$, and $I_{1,3}$ of the input linguistic variable "Age," and the input-value sets $I_{2,1}$, $I_{2,2}$, and $I_{2,3}$ of the input linguistic variable "Property" are obtained as follows:

$$\begin{aligned} O_1 &= \{2000, 2100, 2200\} \\ O_2 &= \{2500, 2600, 2700\} \\ O_3 &= \{3200, 3300\}, I_{1,1} = \{20, 25, 30\} \\ I_{1,2} &= \{45, 50, 60\}, I_{1,3} = \{80, 80\} \\ I_{2,1} &= \{30, 30, 10\}, I_{2,2} = \{50, 30, 10\} \\ I_{2,3} &= \{30, 40\}. \end{aligned}$$

[Step 4]: /* Derive the membership functions of the output fuzzy sets of the output linguistic variable from different α-cuts of the output fuzzy sets corresponding to the output-value sets */

Assume that the threshold value of the α-cut $A_{i,\alpha}(y)$ of the output fuzzy set A_i of the output linguistic variable "Insurance Fee" corresponding to the output-value set O_i of the output linguistic variable "Insurance Fee" equals to 0.8, where $1 \le i \le 3$. Then, based on (9)–(11), we can obtain the membership function of the output fuzzy set A_i of the output linguistic variable "Insurance Fee" represented by the triplet (a_i, b_i, c_i), where $1 \le i \le 3$. The triplet (a_1, b_1, c_1) of the output fuzzy set A_1 of the output linguistic variable "Insurance Fee" can be obtained as follows:

$$\begin{aligned} b_1 &= \frac{2000 + 2200}{2} = 2100 \\ a_1 &= 2100 - \frac{2100 - 2000}{1 - 0.8} = 1600 \\ c_1 &= 2100 + \frac{2200 - 2100}{1 - 0.8} = 2600. \end{aligned}$$

Thus, the membership function $\mu_{A_1}(y)$ of the output fuzzy set A_1 can be represented by the triplet (a_1, b_1, c_1) as follows:

$$\mu_{A_1}(y) = \begin{cases} \dfrac{y - 1600}{500}, & \text{if } 1600 \le y \le 2100, \\ \dfrac{2600 - y}{500}, & \text{if } 2100 \le y \le 2600, \\ 0, & \text{otherwise.} \end{cases}$$

The triplet (a_2, b_2, c_2) of the output fuzzy set A_2 of the output linguistic variable "Insurance Fee" can be obtained as follows:

$$\begin{aligned} b_2 &= \frac{2500 + 2700}{2} = 2600 \\ a_2 &= 2600 - \frac{2600 - 2500}{1 - 0.8} = 2100 \\ c_2 &= 2600 + \frac{2700 - 2600}{1 - 0.8} = 3100. \end{aligned}$$

Thus, the membership function $\mu_{A_2}(y)$ of the output fuzzy set A_2 can be represented by the triplet (a_2, b_2, c_2) as follows:

$$\mu_{A_2}(y) = \begin{cases} \dfrac{y - 2100}{500}, & \text{if } 2100 \le y \le 2600, \\ \dfrac{3100 - y}{500}, & \text{if } 2600 \le y \le 3100, \\ 0, & \text{otherwise.} \end{cases}$$

The triplet (a_3, b_3, c_3) of the output fuzzy set A_3 of the output linguistic variable "Insurance Fee" is obtained as follows:

$$\begin{aligned} b_3 &= \frac{3200 + 3300}{2} = 3250 \\ a_3 &= 3250 - \frac{3250 - 3200}{1 - 0.8} = 3000 \\ c_3 &= 3250 + \frac{3300 - 3250}{1 - 0.8} = 3500. \end{aligned}$$

Thus, the membership function $\mu_{A_3}(y)$ of the output fuzzy set A_3 is represented by the triplet (a_3, b_3, c_3) as follows:

$$\mu_{A_3}(y) = \begin{cases} \dfrac{y - 3000}{250}, & \text{if } 3000 \le y \le 3250, \\ \dfrac{3500 - y}{250}, & \text{if } 3250 \le y \le 3500, \\ 0, & \text{otherwise.} \end{cases}$$

The membership functions of the output fuzzy sets A_1, A_2, and A_3 of the output linguistic variable "Insurance Fee" are shown in Fig. 8.

[Step 5]: /* Sort the input values of the input-value sets of the input linguistic variable */

The two input linguistic variables "Age" and "Property" are denoted by X_1 and X_2. In the input-value sets $I_{1,1}$, $I_{1,2}$, and $I_{1,3}$ of the input linguistic variable "Age," the sorted input-value sets $I'_{1,1}$, $I'_{1,2}$, and $I'_{1,3}$ are the same as the input-value sets $I_{1,1}$, $I_{1,2}$, and $I_{1,3}$, respectively. In the input-value sets $I_{2,1}$, $I_{2,2}$, and $I_{2,3}$ of the input

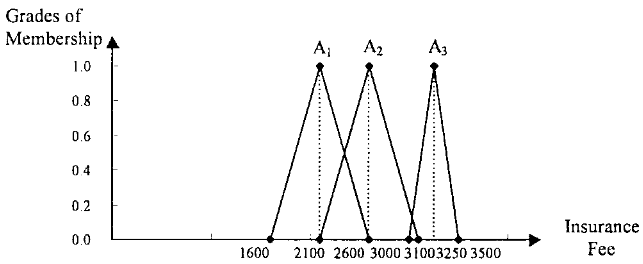

Fig. 8. The fuzzy sets of the output linguistic variable "Insurance Fee."

linguistic variable "Property," the sorted input-value sets are represented by $I'_{2,1}$, $I'_{2,2}$, and $I'_{2,3}$. That is

$$I'_{1,1} = \{20, 25, 30\},\ I'_{1,2} = \{45, 50, 60\}$$
$$I'_{1,3} = \{80, 80\},\ I'_{2,1} = \{10, 30, 30\}$$
$$I'_{2,2} = \{10, 30, 50\},\ I'_{2,3} = \{30, 40\}.$$

[Step 6]: /* Construct the equivalence relation between the input values of the input-value set of the input linguistic variable */

Based on (13), the compatibility relations $R_{1,1}$, $R_{1,2}$, and $R_{1,3}$ are derived from the input-value sets $I_{1,1}$, $I_{1,2}$, and $I_{1,3}$ of the input linguistic variable "Age," respectively,

$$R_{1,1} = \begin{bmatrix} 1 & 0.33 & 0 \\ 0.33 & 1 & 0.33 \\ 0 & 0.33 & 1 \end{bmatrix}$$
$$R_{1,2} = \begin{bmatrix} 1 & 0.6 & 0 \\ 0.6 & 1 & 0.2 \\ 0 & 0.2 & 1 \end{bmatrix}$$
$$R_{1,3} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

where

$$\delta_{1,1} = \frac{(30-20)+(30-25)}{2} = 7.5$$
$$\delta_{1,2} = \frac{(60-45)+(60-50)}{2} = 12.5$$
$$\delta_{1,3} = \text{Inf}$$

where Inf means positive infinity.

The equivalence relations $R^T_{1,1}$, $R^T_{1,2}$, and $R^T_{1,3}$ of the input-value sets $I_{1,1}$, $I_{1,2}$, and $I_{1,3}$ of the input linguistic variable "Age" can be obtained by the max–min transitive closure of compatibility relations $R_{1,1}$, $R_{1,2}$, and $R_{1,3}$, respectively

$$R^T_{1,1} = \begin{bmatrix} 1 & 0.33 & 0.33 \\ 0.33 & 1 & 0.33 \\ 0.33 & 0.33 & 1 \end{bmatrix},$$
$$R^T_{1,2} = \begin{bmatrix} 1 & 0.6 & 0.2 \\ 0.6 & 1 & 0.2 \\ 0.2 & 0.2 & 1 \end{bmatrix},$$
$$R^T_{1,3} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

where

$$\delta_{2,1} = \frac{(30-10)+(30-30)}{2} = 10$$
$$\delta_{2,2} = \frac{(50-10)+(50-30)}{2} = 30$$
$$\delta_{2,3} = 40 - 30 = 10.$$

By the same way, the compatibility relations of the input-value sets $I'_{2,1}$, $I'_{2,2}$, and $I'_{2,3}$ of the input linguistic variable "Property" can be obtained as follows:

$$R_{2,1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$
$$R_{2,2} = \begin{bmatrix} 1 & 0.33 & 0 \\ 0.33 & 1 & 0.33 \\ 0 & 0.33 & 1 \end{bmatrix}$$
$$R_{2,3} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then, the equivalence relations of the sorted input-value sets $I'_{2,1}$, $I'_{2,2}$, and $I'_{2,3}$ of the input linguistic variable "Property" can be obtained as follows:

$$R^T_{2,1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$
$$R^T_{2,2} = \begin{bmatrix} 1 & 0.33 & 0.33 \\ 0.33 & 1 & 0.33 \\ 0.33 & 0.33 & 1 \end{bmatrix}$$
$$R^T_{2,3} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

[Step 7]: /* Divide the input-value set into different input-value subsets based on the α-cuts of the equivalence relation between the input values of the input-value set */

Indeed, the training examples are so few that it is not necessary to partition input values again. The input-value subsets of the input linguistic variables "Age" and "Property" are the same with their corresponding input-value sets, and the number of subsets for each input-value set equals to one. $T_1(X_1) = T_2(X_1) = T_3(X_1) = 1$, $T_1(X_2) = T_2(X_2) = T_3(X_2) = 1$. Hence, the input-value subsets of the input linguistic variable "Age" are obtained as follows:

$$I_{1,1,1} = I'_{1,1} = \{20, 25, 30\}$$
$$I_{1,2,1} = I'_{1,2} = \{45, 50, 60\}$$
$$I_{1,3,1} = I'_{1,3} = \{80, 80\}.$$

The input-value subsets of the input linguistic variable "Property" are obtained as follows:

$$I_{2,1,1} = I'_{2,1} = \{10, 30, 30\},$$
$$I_{2,2,1} = I'_{2,2} = \{10, 30, 50\},$$
$$I_{2,3,1} = I'_{2,3} = \{30, 40\}.$$

[Step 8]: /* Derive the membership functions of the input fuzzy sets of the input linguistic variable from the α-cut of the input fuzzy set corresponding to the input-value subset */

The input-value subsets $I_{1,i,j}$ can be thought of as the α-cut set $A_{(1,i,j),\alpha}$ of the input fuzzy set $A_{1,i,j}$ of the input linguistic variable "Age," where $i = 1, 2, 3$, and $j = 1$. Assume that the threshold value of the α-cut set $A_{(1,i,j),\alpha}$ is equal to 0.7. Thus, we can derive the membership function of the input fuzzy set $A_{1,i,j}$ of the input linguistic variable "Age" based on (17)–(19). The triplet $(a_{1,1,1}, b_{1,1,1}, c_{1,1,1})$ of the input fuzzy set $A_{1,1,1}$ of the input linguistic variable "Age" can be obtained as follows:

$$(a_{1,1,1}, b_{1,1,1}, c_{1,1,1}) = (8.33, 25, 41.67).$$

The membership function $\mu_{A_{1,1,1}}(x_1)$ of the input fuzzy set $A_{1,1,1}$ of the input linguistic variable "Age" can be represented by the triplet $(a_{1,1,1}, b_{1,1,1}, c_{1,1,1})$ as follows:

$$\mu_{\mu_{A_{1,1,1}}}(x_1) = \begin{cases} \dfrac{x_1 - 8.33}{16.67}, & \text{if } 8.33 \le x_1 \le 25, \\ \dfrac{41.67 - x_1}{16.67}, & \text{if } 25 \le x_1 \le 41.67, \\ 0, & \text{otherwise.} \end{cases}$$

The triplet $(a_{1,2,1}, b_{1,2,1}, c_{1,2,1})$ of the input fuzzy set $A_{1,2,1}$ of the input linguistic variable "Age" can be obtained as follows:

$$(a_{1,2,1}, b_{1,2,1}, c_{1,2,1}) = (27.5, 52.5, 77.5).$$

The membership function $\mu_{A_{1,2,1}}(x_1)$ of the input fuzzy set $A_{1,2,1}$ of the input linguistic variable "Age" can be represented by the triplet $(a_{1,2,1}, b_{1,2,1}, c_{1,2,1})$ as follows:

$$\mu_{A_{1,2,1}}(x_1) = \begin{cases} \dfrac{x_1 - 27.5}{25}, & \text{if } 27.5 \le x_1 \le 52.5, \\ \dfrac{77.5 - x_1}{25}, & \text{if } 52.5 \le x_1 \le 77.5, \\ 0, & \text{otherwise.} \end{cases}$$

The triplet $(a_{1,3,1}, b_{1,3,1}, c_{1,3,1})$ of the input fuzzy set $A_{1,3,1}$ of the input linguistic variable "Age" can be obtained as follows:

$$(a_{1,3,1}, b_{1,3,1}, c_{1,3,1}) = (36.67, 70, 103.33)$$

where the value of $b_{1,3,1}$ is equal to the average of the larget element of $I_{1,2,1}$ and the smallest element of $I_{1,3,1}$, i.e.,

$$b_{1,3,1} = \frac{\text{Max}(40, 50, 60) + \text{Min}(80, 80)}{2} = 70.$$

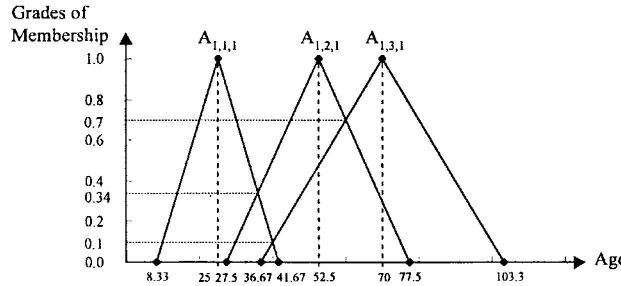

Fig. 9. The fuzzy sets of the input linguistic variable "Age."

The membership function $\mu_{A_{1,3,1}}(x_1)$ of the input fuzzy set $A_{1,3,1}$ of the input linguistic variable "Age" can be represented by the triplet $(a_{1,3,1}, b_{1,3,1}, c_{1,3,1})$ as follows:

$$\mu_{A_{1,3,1}}(x_1) = \begin{cases} \dfrac{x_1 - 36.67}{33.33}, & \text{if } 36.67 \le x_1 \le 70, \\ \dfrac{103.33 - x_1}{33.33}, & \text{if } 70 \le x_1 \le 103.33, \\ 0, & \text{otherwise.} \end{cases}$$

The input fuzzy sets $A_{1,1,1}$, $A_{1,2,1}$, and $A_{1,3,1}$ of the input linguistic variable "Age" are shown in Fig. 9. The membership functions $\mu_{A_{2,1\,1}}(x_2)$, $\mu_{A_{2,2,1}}(x_2)$, and $\mu_{A_{2,3,1}}(x_2)$ of the input fuzzy sets $A_{2,1,1}$, $A_{2,2,1}$, $A_{2,3,1}$ of the input linguistic variable "Property" can be obtained as follows:

$$\mu_{A_{2,1,1}}(x_2) = \begin{cases} \dfrac{x_2}{20}, & \text{if } 0 \le x_2 \le 20, \\ \dfrac{53.33 - x_2}{33.33}, & \text{if } 20 \le x_2 \le 53.33, \\ 0, & \text{otherwise.} \end{cases}$$

$$\mu_{A_{2,2,1}}(x_2) = \begin{cases} \dfrac{x_2}{30}, & \text{if } 0 \le x_2 \le 30, \\ \dfrac{96.67 - x_2}{66.67}, & \text{if } 30 \le x_2 \le 96.67, \\ 0, & \text{otherwise.} \end{cases}$$

$$\mu_{A_{2,3,1}}(x_2) = \begin{cases} \dfrac{x_2 - 18.33}{16.67}, & \text{if } 18.33 \le x_2 \le 35, \\ \dfrac{51.67 - x_2}{16.67}, & \text{if } 35 \le x_2 \le 51.67, \\ 0, & \text{otherwise.} \end{cases}$$

The input fuzzy sets $A_{2,1,1}$, $A_{2,2,1}$, and $A_{2,3,1}$ of the input linguistic variable "Property" are shown in Fig. 10.

[Step 9]: /* Construct fuzzy rules based on the hierarchical relationships between the output fuzzy sets of the output linguistic variable and the corresponding input fuzzy sets of the input linguistic variables */

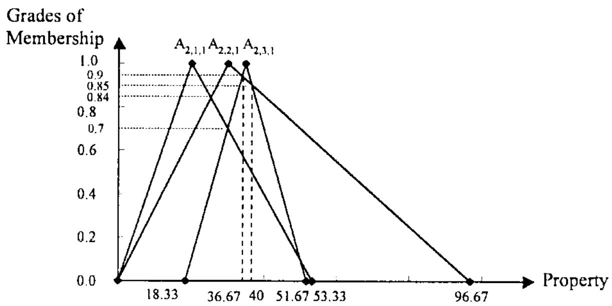

Fig. 10. The fuzzy sets of the input linguistic variable "Property."

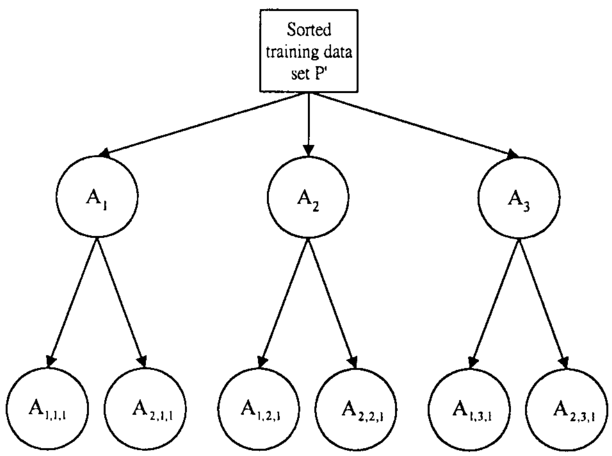

Fig. 11. The hierarchical relationship between the output fuzzy set A_j of the output linguistic variable "Insurance Fee," the input fuzzy set $A_{1,\,j,\,k_1}$ of the input linguistic variable "Age," and the input fuzzy set $A_{2,\,j,\,k_2}$ of the input linguistic variable "Property."

The hierarchical relationship between the output fuzzy set A_j of the output linguistic variable "Insurance Fee," the input fuzzy set $A_{1,\,j,\,k_1}$ of the input linguistic variable "Age," and the input fuzzy set $A_{2,\,j,\,k_2}$ of the input linguistic variable "Property" is shown in Fig. 11.

Based on the hierarchical relationship between the output fuzzy set A_j of the output linguistic variable "Insurance Fee," the input fuzzy set $A_{1,\,j,\,k_1}$ of the input linguistic variable "Age," and the input fuzzy set $A_{2,\,j,\,k_2}$ of the input linguistic variable "Property," the fuzzy rules of the Insurance Fee problem are generated as follows:

IF Age is $A_{1,\,1,\,1}$ **AND** Property is $A_{2,\,1,\,1}$
THEN Insurance Fee is A_1
IF Age is $A_{1,\,2,\,1}$ **AND** Property is $A_{2,\,2,\,1}$
THEN Insurance Fee is A_2
IF Age is $A_{1,\,3,\,1}$ **AND** Property is $A_{2,\,3,\,1}$
THEN Insurance Fee is A_3.

[Step 10]: /* Calculate the equality degree between the input fuzzy sets of the input linguistic variable */

Based on (25), we can calculate the equality degree between the input fuzzy sets $A_{1,\,1,\,1}$, $A_{1,\,2,\,1}$, $A_{1,\,3,\,1}$ of the input linguistic variable "Age" as follow.

First, the area of the intersection of the two distinct input fuzzy sets $A_{1,\,1,\,1}$ and $A_{1,\,2,\,1}$ can be calculated from Fig. 9 as follows:

$$|A_{1,\,1,\,1} \cap A_{1,\,2,\,1}| = \frac{(41.67 - 27.5) \times 0.34}{2} = 2.4089.$$

The area of the union of the two distinct input fuzzy sets $A_{1,\,1,\,1}$ and $A_{1,\,2,\,1}$ can be calculated as from Fig. 9 as shown in (33) at the bottom of the page. Based on (25), the equality degree between fuzzy sets $A_{1,\,1,\,1}$ and $A_{1,\,2,\,1}$ is obtained as follows:

$$E(A_{1,\,1,\,1},\, A_{1,\,2,\,1}) = \frac{|A_{1,\,1,\,1} \cap A_{1,\,2,\,1}|}{|A_{1,\,1,\,1} \cup A_{1,\,2,\,1}|} = \frac{2.4089}{39.2611} = 0.0613.$$

By the same way, the equality degrees, $E(A_{1,\,1,\,1},\, A_{1,\,3,\,1})$ and $E(A_{1,\,2,\,1},\, A_{1,\,3,\,1})$, are calculated as shown in (34) at the bottom of the next page. The equality matrix E_{Age} of the input fuzzy sets of the input linguistic variable "Age" is obtained as follows:

$$E_{\text{Age}} = \begin{array}{c} \\ A_{1,1,1} \\ A_{1,2,1} \\ A_{1,3,1} \end{array} \begin{array}{c} \begin{array}{ccc} A_{1,1,1} & A_{1,2,1} & A_{1,3,1} \end{array} \\ \begin{pmatrix} 1 & 0.0613 & 0.005 \\ 0.0613 & 1 & 0.3246 \\ 0.005 & 0.3246 & 1 \end{pmatrix} \end{array}.$$

Based on (25), we can calculate the equality degree between the input fuzzy sets $A_{2,\,1,\,1}$, $A_{2,\,2,\,1}$, $A_{2,\,3,\,1}$ of the input linguistic variable "Property" as shown in (35) at the bottom of the next page. The equality matrix E_{Property} of the input fuzzy sets of the input linguistic variable "Property" is obtained as follows:

$$E_{\text{Property}} = \begin{array}{c} \\ A_{2,1,1} \\ A_{2,2,1} \\ A_{2,3,1} \end{array} \begin{array}{c} \begin{array}{ccc} A_{2,1,1} & A_{2,2,1} & A_{2,3,1} \end{array} \\ \begin{pmatrix} 1 & 0.4258 & 0.3685 \\ 0.4258 & 1 & 0.41 \\ 0.3685 & 0.41 & 1 \end{pmatrix} \end{array}$$

Assume that the value of α_{Equality} specified by the user is 0.5, then we can see that the equality degrees between the input fuzzy sets of the

$$|A_{1,\,1,\,1} \cup A_{1,\,2,\,1}| = \frac{(41.67 - 8.33) \times 1 + (77.5 - 27.5) \times 1 - (41.67 - 27.5) \times 0.34}{2} = 39.2611 \qquad (33)$$

input linguistic variables "Age" and "Property" in the equality matrices E_{Age} and E_{Property}, respectively, are less than 0.5. It is not necessary to perform the merge operation.

Finally, the fuzzy rules generated by the proposed fuzzy learning algorithm are listed as follows:

IF Age is $A_{1,1,1}$ **AND** Property is $A_{2,1,1}$
THEN Insurance Fee is A_1
IF Age is $A_{1,2,1}$ **AND** Property is $A_{2,2,1}$
THEN Insurance Fee is A_2
IF Age is $A_{1,3,1}$ **AND** Property is $A_{2,3,1}$
THEN Insurance Fee is A_3.

V. Experimental Results

There are three kinds of flowers in the Iris data [7], i.e., Setosa, Versicolor, and Verginica, where each flower can be identified by the four kinds of input variables, i.e., sepal length, sepal width, petal length, and petal width. The unit of each input variable is centimeters. The Iris data set contains 150 data. Based on the proposed fuzzy learning algorithm, we have implemented a program on a Pentium PC to generate the membership functions and the fuzzy rules from the Iris training data automatically and then evaluate the average accuracy rate of the proposed fuzzy learning algorithm with the testing data set. We use the MATLAB software to develop the program and randomly choose 50% of the Iris data as the training data set and the other 50% as the testing data set. The membership functions of the input variables are derived under the 0.5-cut of the equality relations and shown in Figs. 12–15. The final generated fuzzy rules are shown in Table II.

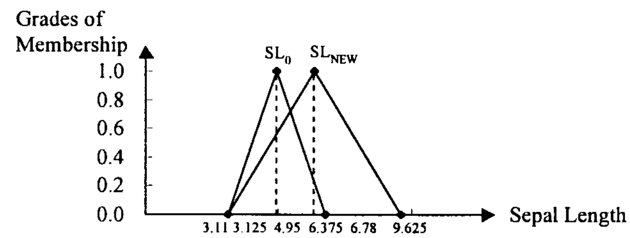

Fig. 12. Fuzzy sets of the input variable "Sepal Length."

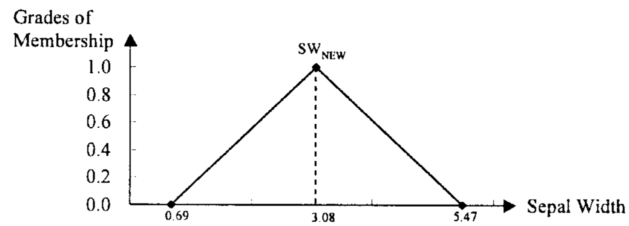

Fig. 13. Fuzzy sets of the input variable "sepal width."

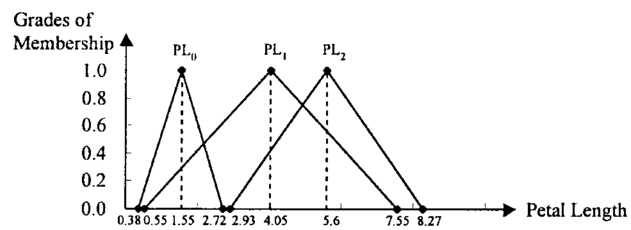

Fig. 14. Fuzzy sets of the input variable "petal length."

The simplified fuzzy rules are as follows.

R_0: **IF** sepal length is SL_0 **AND** sepal width is SW_{NEW} **AND** petal length is PL_0 **AND** petal width is PW_0 **THEN** the flower is Setosa.

$$\begin{aligned} E(A_{1,1,1}, A_{1,3,1}) &= \frac{|A_{1,1,1} \cap A_{1,3,1}|}{|A_{1,1,1} \cup A_{1,3,1}|} = \frac{(41.67 - 36.67) \times 0.1}{(41.67 - 8.33) \times 1 + (103.3 - 36.67) \times 1 - (41.67 - 36.67) \times 0.1} \\ &= 0.005, \\ E(A_{1,2,1}, A_{1,3,1}) &= \frac{|A_{1,2,1} \cap A_{1,3,1}|}{|A_{1,2,1} \cup A_{1,3,1}|} = \frac{(77.5 - 36.67) \times 0.7}{(77.5 - 27.5) \times 1 + (103.3 - 36.67) \times 1 - (77.5 - 36.67) \times 0.7} \\ &= 0.3246 \end{aligned} \tag{34}$$

$$\begin{aligned} &E(A_{2,1,1}, A_{2,2,1}) \\ &\quad = \frac{|A_{2,1,1} \cap A_{2,2,1}|}{|A_{2,1,1} \cup A_{2,2,1}|} = \frac{53.33 \times 0.84}{53.33 + 96.67 - 53.33 \times 0.84} = 0.4258, \\ &E(A_{2,1,1}, A_{2,3,1}) \\ &\quad = \frac{|A_{2,1,1} \cap A_{2,3,1}|}{|A_{2,1,1} \cup A_{2,3,1}|} = \frac{(51.67 - 18.33) \times 0.7}{53.33 + 33.34 - (51.67 - 18.33) \times 0.7} = 0.3685, \\ &E(A_{2,2,1}, A_{2,3,1}) \\ &\quad = \frac{|A_{2,2,1} \cap A_{2,3,1}|}{|A_{2,2,1} \cup A_{2,3,1}|} \\ &\quad = \frac{(40 - 18.33) \times 0.85 + (0.85 + 0.9) \times (40 - 36.67) + (51.67 - 36.67) \times 0.9}{(51.67 - 18.33) \times 1 + (96.67 - 0) \times 1 - [(40 - 18.33) \times 0.85 + (0.85 + 0.9) \times (40 - 36.67) + (51.67 - 36.67) \times 0.9]} \\ &\quad = 0.41 \end{aligned} \tag{35}$$

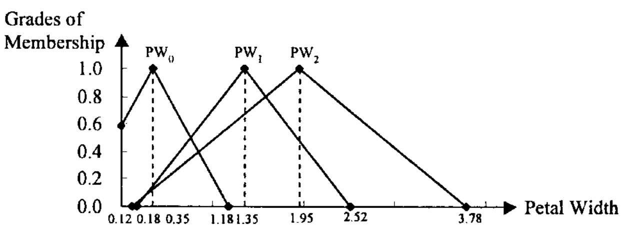

Fig. 15. Fuzzy sets of the input variable "petal width."

TABLE II
SIMPLIFIED FUZZY CLASSIFICATION RULES FOR THE IRIS DATA

Sepal Length	Sepal Width	Petal Length	Petal Width	Result
SL_0	SW_{new}	PL_0	PW_0	Setosa
SL_{new}	SW_{new}	PL_1	PW_1	Versicolor
SL_{new}	SW_{new}	PL_2	PW_2	Virginica

TABLE III
THE AVERAGE ACCURACY RATE OF THE PROPOSED FUZZY ALGORITHM FOR THE IRIS DATA

Setosa	Versicolor	Virginica	Average
100%	93.38%	95.24%	96.21%

TABLE IV
A COMPARISON OF THE NUMBER OF FUZZY RULES AND THE NUMBER OF INPUT FUZZY SETS BETWEEN HONG-AND-LEE'S ALGORITHM AND THE PROPOSED ALGORITHM

	Hong-and-Lee's Algorithm [11]	The Proposed Algorithm
Number of Rules	6.21	3
Number of Input Fuzzy Sets	8	8.21

R_1: **IF** sepal length is SL_{NEW} **AND** sepal width is SW_{NEW} **AND** petal length is PL_1 **AND** petal width is PW_1 **THEN** the flower is Versicolor.

R_2: **IF** sepal length is SL_{NEW} **AND** sepal width is SW_{NEW} **AND** petal length is PL_2 **AND** petal width is PW_2 **THEN** the flower is Virginica.

The average accuracy rate of the proposed fuzzy learning algorithm after 200 runs is listed in Table III.

A comparison of the proposed fuzzy learning algorithm and Hong and Lee's algorithm [11] in terms of the number of the fuzzy rules and the number of the input fuzzy sets are shown in Table IV.

A comparison of the experiment results between Hong and Lee's algorithm [11] and the proposed algorithm is listed in Table V.

From Tables IV and V, we can see that the average accuracy rate of the proposed algorithm is better than that of Hong and Lee's algorithm. The number of rules generated by the proposed algorithm is less than the number of rules generated by Hong and Lee's algorithm. Furthermore, the proposed algorithm does not need to predefine any membership functions of the input variables and the output variables. The membership functions and fuzzy rules are generated by the proposed fuzzy learning algorithm from the numerical training data.

TABLE V
A COMPARISON OF THE AVERAGE ACCURACY RATE BETWEEN HONG-AND-LEE'S ALGORITHM AND THE PROPOSED ALGORITHM

	Hong-and-Lee's Algorithm [11]	The Proposed Algorithm
Average Accuracy Rate	95.57%	96.21%

VI. CONCLUSIONS

In this paper, we have presented a new fuzzy learning algorithm based on the α-cuts of equivalence relations and the α-cuts of fuzzy sets to construct membership functions and to generate fuzzy rules from numerical training data. Furthermore, we also apply the proposed algorithm to the Iris data classification problem. Based on the proposed algorithm, we have implemented a program on a Pentium PC using MATLAB Version 4.0 to deal with the Iris data classification problem. The experimental results are compared with the results of Hong and Lee's learning algorithm [11]. The proposed algorithm is better than the one presented in [11] due to the following facts.

1) The proposed algorithm could get a better average accuracy rate than the one presented in [11]. From the experimental results shown in the previous section, we can see that the average accuracy rate of the proposed algorithm is 96.21%, where the average accuracy rate of Hong and Lee's algorithm is 95.57%.
2) The proposed algorithm generates fewer fuzzy rules than the one proposed in [11]. From the experimental results shown in Section V, we can see that the number of fuzzy rules generated by the proposed algorithm is 3, but the number of fuzzy rules generated by Hong and Lee's algorithm is 6.21.
3) We do not need to partition the input spaces and the output spaces into fuzzy regions or predefine any membership functions as shown in [11]. The membership functions and fuzzy rules can be automatically derived from the numerical training data by the proposed algorithm.

REFERENCES

[1] G. Beliakov, "Fuzzy sets and membership functions based on probability," *Inf. Sci.*, vol. 91, no. 1, pp. 95–111, 1996.
[2] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms.* New York: Plenum, 1981, pp. 65–86.
[3] C. T. Chao, Y. J. Chen, and C. C. Teng, "Simplification of fuzzy-neural systems using similarity analysis," *IEEE Trans. Syst., Man, Cybern. B*, vol. 26, pp. 344–354, Apr. 1996.
[4] S. M. Chen, M. S. Yeh, and P. Y. Hsiao, "A comparison of similarity measures of fuzzy values," *Fuzzy Sets Syst.*, vol. 72, no. 1, pp. 79–89, 1995.

[5] S. M. Chen, "Using fuzzy reasoning techniques for fault diagnosis of J-85 jet engines," in *Proc. 3rd Nat. Conf. Science Technology National Defense,* Taoyuan, Taiwan, R.O.C., 1994, vol. 1, pp. 29–34.
[6] D. Dubois and H. Prade, "What are fuzzy rules and how to use them," *Fuzzy Sets Syst.,* vol. 84, no. 2, pp. 169–185, 1996.
[7] R. Fisher, "The use of multiple measurements in taxonomic problems," *Ann. Eugenics,* vol. 7, pp. 179–188, 1936.
[8] J. Giarratano and G. Riley, *Expert Systems: Principles and Programming.* Boston, MA: PWS, 1994.
[9] A. Grauel and H. Mackenberg, "Mathematical analysis of the Sugeno controller leading to general design rules," *Fuzzy Sets Syst.,* vol. 85, no. 2, pp. 165–175, 1997.
[10] K. Hattori and Y. Tor, "Effective algorithms for the nearest neighbor method in the clustering problem," *Pattern Recognit.,* vol. 26, no. 5, pp. 741–746, 1993.
[11] T. P. Hong and C. Y. Lee, "Induction of fuzzy rules and membership functions from training examples," *Fuzzy Sets Syst.,* vol. 84, no. 1, pp. 33–47, 1996.
[12] T. Ichimura and E. Tazaki, "Learning and extracting method for fuzzy rules using neural networks with modified structure level adaptation," in *Proc. 3rd Int. Conf. Fuzzy Logic, Neural Nets, Soft Computing,* Japan, 1994, pp. 87–90.
[13] H. Ishibuchi and H. Tanaka, "Neural network that learn from fuzzy if–then rules," *IEEE Trans. Fuzzy Syst.,* vol. 1, pp. 85–97, 1993.
[14] J. Rudin, *Principles of Mathematical Analysis.* New York: McGraw-Hill, 1976.
[15] N. K. Kasabov, "Learning fuzzy rules and approximate reasoning in fuzzy neural networks and hybrid systems," *Fuzzy Sets Syst.,* vol. 82, no. 2, pp. 135–149, 1996.
[16] F. Klawonn and R. Kruse, "Constructing a fuzzy controller from data," *Fuzzy Sets Syst.,* vol. 85, no. 2, pp. 177–193, 1997.
[17] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic Theory and Applications.* Englewood Cliffs, NJ: Prentice-Hall, 1995.
[18] Y. Kodratoff and R. S. Michalski, *Machine Learning: An Artificial Intelligence Approach.* San Mateo: Morgan Kaufmann, 1990, vol. 3, pp. 113–127.
[19] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence.* Englewood Cliffs: Prentice-Hall, 1991, pp. 299–335.
[20] R. Langari and L. Wang, "Fuzzy models, modular networks, and hybrid learning," *Fuzzy Sets Syst.,* vol. 79, no. 2, pp. 141–150, 1996.
[21] C. T. Lin and C. S. George Lee, "Neural-network-based fuzzy logic control and decision system," *IEEE Trans. Comput.,* vol. 40, pp. 1320–1336, Dec. 1991.
[22] ———, *Neural Fuzzy Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1996.
[23] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, *Machine Learning: An Artificial Intelligence Approach.* Los Altos, CA: Morgan Kaufmann, 1983, vol. 1, p. 41–80.
[24] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, *Machine Learning: An Artificial Intelligence Approach.* Los Altos, CA: Morgan Kaufmann, 1983, vol. 2, pp. 45–61.
[25] R. S. Michalski and G. Tecuci, *Machine Learning: A Multistrategy Approach.* San Francisco, CA: Morgan Kaufmann, 1994, vol. 4.
[26] K. Nozaki, H. Ishibuchi, and H. Tanaka, "A simple but powerful heuristic method for generating fuzzy rules from numerical data," *Fuzzy Sets Syst.,* vol. 86, no. 3, pp. 251–270, 1997.
[27] D. Park, A. Kandel, and G. Langholz, "Genetic-based new fuzzy reasoning models with application to fuzzy control," *IEEE Trans. Syst., Man, Cybern.,* vol. 24, pp. 39–47, Jan. 1994.
[28] P. K. Simpson, "Fuzzy min–max neural networks—Part 2: Clustering," *IEEE Trans. Fuzzy Syst.,* vol. 1, pp. 32–45, 1993.
[29] J. J. Shann and H. C. Fu, "A fuzzy neural network for rule acquiring on fuzzy control systems," *Fuzzy Sets Syst.,* vol. 71, no. 3, pp. 345–357, 1995.
[30] K. Uehara and M. Fujise, "Fuzzy inference based on families of α-level sets," *IEEE Trans. Fuzzy Syst.,* vol. 1, pp. 111–124, 1993.
[31] W. J. Wang, "New similarity measures on fuzzy sets and on elements," *Fuzzy Sets Syst.,* vol. 85, no. 3, pp. 305–309, 1997.
[32] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Trans. Syst., Man, Cybern.,* vol. 22, pp. 1414–1427, 1992.
[33] L. X. Wang, *Adaptive Fuzzy Systems and Control: Design and Stability Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1994, pp. 27–28.
[34] T. P. Wu and S. M. Chen, "A new method for constructing membership functions and fuzzy rules from numerical data," in *Proc. 1997 5th Nat. Conf. Fuzzy Theory Applications,* Tainan, Taiwan, R.O.C., 1997, pp. 499–504.
[35] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst., Man, Cybern.,* vol. 15, pp. 116–132, 1985.
[36] R. R. Yager and D. P. Filev, *Essential Fuzzy Modeling and Control.* New York: Wiley, 1994.
[37] V. R. Young, "Fuzzy subsethood," *Fuzzy Sets Syst.,* vol. 77, no. 3, pp. 371–384, 1996.
[38] L. A. Zadeh, "Fuzzy sets," *Inf. Contr.,* vol. 8, pp. 338–353, 1965.
[39] ———, "Fuzzy logic," *IEEE Comput.,* vol. 21, pp. 83–91, 1988.
[40] ———, "The concepts of a linguistic variable and its application to approximate reasoning (I)," *Inf. Sci.,* vol. 8, pp. 199–249, 1975.
[41] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications.* Boston, MA: Kluwer-Nijhoff, 1991.

Tzu-Ping Wu was born on October 9, 1967, in Kaohsiung, Taiwan, R.O.C. He received the B.S. degree in information engineering from the Feng-Chia University, Taichung, Taiwan, in June 1989, and the M.S. degree in computer and information science from the National Chiao Tung University, Hsinchu, Taiwan, in June 1997.

He is currently with the Chung Shan Institute of Science and Technology, Taoyuan, Taiwan. His current research interests include fuzzy systems, expert systems, and artificial intelligence.

Shyi-Ming Chen (M'88–SM'96) was born on January 16, 1960, in Taipei, Taiwan, R.O.C. He received the B.S. degree in electronic engineering from the National Taiwan University of Science and Technology, Taipei, in June 1982, and the M.S. and Ph.D. degrees in electrical engineering from the National Taiwan University, Taipei, in June 1986 and June 1991, respectively.

From August 1987 to July 1989 and from August 1990 to July 1991, he was with the Department of Electronic Engineering, Fu-Jen University, Taipei. From August 1991 to July 1996, he was an Associate Professor in the Department of Computer and Information Science, National Chiao Tung University, Hsinchu, Taiwan. From August 1996 to July 1998, he was a Professor in the Department of Computer and Information Science, National Chiao Tung University, Hsinchu. Since August 1998, he has been a Professor in the Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei. His current research interests include fuzzy systems, database systems, expert systems, and artificial intelligence. He has published more than 95 papers in referred journals and conference proceedings.

Dr. Chen was the winner of the 1994 Outstanding Paper Award of the *Journal of Computer and Information Education* and the winner of the 1995 Outstanding Paper Award of the Computer Society of the Republic of China. He was the winner of the 1997 Outstanding Youth Electrical Engineer Award of the Chinese Institute of Electrical Engineering, Republic of China. He is a member of the ACM, the International Fuzzy Systems Association (IFSA), and the Phi Tau Phi Scholastic Honor Society. He is an Executive Committee Member of the Taiwanese Association for Artificial Intelligence (TAAI). He is also an Executive Committee Member of the Chinese Fuzzy Systems Association (CFSA). He is an Editor of the *Journal of the Chinese Grey Systems Association*. He is listed in *International Who's Who of Professionals*. He is also listed in Marquis' *Who's Who in the World* (16th ed.).