CrossMark

Mining Fuzzy Association Rules in the Framework of AFS Theory

Bo Wang[1] · **Xiao-dong Liu**[1] · **Li-dong Wang**[1]

Abstract In this paper, firstly we study the representations and fuzzy logic operations for the fuzzy concepts in real data systems. Secondly, we propose a new fuzzy association rule mining algorithm in the framework of AFS (Axiomatic Fuzzy Sets) theory. Compared with the current algorithms, the advantage of proposed algorithm has two advantages. One is that the membership functions of the fuzzy sets representing the extracted rules and the fuzzy logic operations applied to extract fuzzy rules are determined by the distribution of the data, instead of the fuzzy sets defined by some special functions, t-norm, t-conorm, negation operator, implication operator and fuzzy similarity relation given in advance. The extracted fuzzy rules are interpretable and similar to human intuition. Another is that its simplicity in implementation and mathematical beauty in fuzzy theory, and can be directly applied to extract fuzzy association rules in real data systems. Finally, a well-known example Iris dataset is used to illustrate the effectiveness of the new algorithm based on the proposed degrees of implication. We obtained reclassification accuracy 98 %.

Keywords Molecular lattices · AFS fuzzy logic · Degrees of implication ·
Fuzzy association rules

✉ Xiao-dong Liu
  xdliuros@hotmail.com

1 Department of Mathematics, Dalian Maritime University, Dalian 116026,
  People's Republic of China

1 Introduction

Currently, the digitized data are easy to capture and cheap to store. However, raw digital data is rarely of useful in practice and the capability to extract information from raw data is extremely important for decision support [1]. Knowledge discovery, whose objective is to obtain useful knowledge from data stored in large repositories, is recognized as a basic necessity in many areas. Data mining is the step in the knowledge discovery process that attempts to discover novel and meaningful patterns in data. On an ordinary crisp set, correlation coefficient is commonly used in conventional statistics [2,3] as a measure of the linear relationship between two attributes. Since Agrawal et al. introduced the notion of association rules in 1993 [4], association rules mining has attracted many research efforts along with a large number of association rule applications in various fields [5]. In [6], the algorithm proposed by Srikant and Agrawal proceeded by partitioning attribute domains into several intervals and transforming quantitative values into binary ones in order to apply the classical mining algorithm.Fuzzy set technology are applied to knowledge discovery by such as fuzzy extensions of association rules and approximate dependencies base on fuzzy sets, fuzzy logic operators, fuzzy implication operators and fuzzy similarity relations [7–12]. Other fuzzy extensions of association rules include, weighted association rules [13] and different fuzziness-related interestingness measures [14]. The mining results based on the above mentioned approaches for a data systems are all strongly dependent on the options of fuzzy logic operators, fuzzy implication operator, the membership functions defined for the fuzzy sets and the fuzzy similarity relation given in advance.

In real-world applications, the intelligent systems are usually very large and complex due to large dimensions and different type, size and distribution of databases. They often involve a large number of concepts such that it is very hard, if not impossible, to properly define all membership functions by human intuition manually. Also it is hard to choose suitable logic operators which can correctly represent the logic relations among the involved fuzzy sets on the database in advance. In [15,16], the authors have proposed and studied AFS theory in which the membership functions and the fuzzy logic operations can be automatically determined by the distribution of the original data.

In this paper, firstly by well-known Iris dataset, we introduce the algorithm for automatically determining membership functions for fuzzy sets according to the original data or training examples and the logic operations of the fuzzy sets implemented by the AFS algebra. And also some basic ideas, definitions and results of AFS theory as preliminary of this paper are illustrated by the examples for Iris dataset. Secondly, by the analysis of current definitions of the degree of implication, we find the conflict between the current definition and human intuition. In order to overcome the drawback, we propose a new degree of implication based on AFS fuzzy logic and give a new rule extraction algorithm based on the proposed degree of implication. The proposed algorithm is similar to human intuition and can be directly applied to real world applications. Finally we apply the proposed rule extraction algorithm to the well-known Iris dataset and compare its result with other rule extraction algorithms. The illustrative examples show that AFS theory offer a far more flexible and powerful framework for

acquisition and representation human knowledge and studying large-scale intelligent systems in real world applications.

2 Fuzzy Concepts Expressed by AFS Algebra over Attributes

The Iris data [17] has a 150×4 matrix $W = (w_{ij})_{150 \times 4}$ evenly distributed in three classes: iris-setosa, iris-versicolor, and iris-virginica. Vector of sample i, $(w_{i1}, w_{i2}, w_{i3}, w_{i4})$ has four features: sepal length and width, and petal length and width (all given in centimeters). Let $I = \{I_1, I_2, I_3, I_4, I_5\}$ be the set of attributes, where I_1 is sepal length, I_2 is sepal width, I_3 is petal length, I_4 is petal width and I_5 is the labels of class 1, 2, 3 (i.e. 1, 2, 3 for class iris-setosa, iris-versicolor, iris-virginica respectively) are the attributes of the original quantitative database $D = \{t_1, t_2, \ldots, t_{150}\}$ for the 150 samples, $\forall i \in [1, 50], t_i[I_5] = 1$, i.e. samples in class 1; $\forall i \in [51, 100], t_i[I_5] = 2$, i.e. samples in class 2; $\forall i \in [101, 150], t_i[I_5] = 3$, i.e. samples in class 3; $t_i[I_j] = \frac{w_{ij}}{\max_j(w_{ij})}$, $\forall i \in [1, 150]$, $j \in [1, 4]$ (i.e. $t_i[I_j]$ is normalized in interval $[0, 1]$). In this paper, we will apply the proposed fuzzy association rule mining algorithm to find fuzzy association rules for class 1, 2, 3. Let $c_{1j} = \frac{\sum_{1 \leq i \leq 50} t_i[I_j]}{50}$, $c_{2j} = \frac{\sum_{51 \leq i \leq 100} t_i[I_j]}{50}$, $c_{3j} = \frac{\sum_{101 \leq i \leq 150} t_i[I_j]}{50}$, $j \in [1, 4]$ i.e. the mean of class 1, 2, 3 on attribute I_j. The following two fuzzy rules are the fuzzy descriptions of "*iris-setosa*" for a classification model.

Rule $\mathcal{R}_1$: If $t[I_1]$ is about c_{11} and $t[I_2]$ is about c_{22}, then t is "*iris-setosa*";
Rule $\mathcal{R}_2$: If $t[I_3]$ is not about c_{23} and $t[I_2]$ is about c_{32} and $t[I_4]$ is not about c_{14}, then t is "*iris-setosa*";

Let $M = \{m_{j,k} | 1 \leq j \leq 4, 1 \leq k \leq 2\}$ be the set of fuzzy terms, where $m_{j,1}, m_{j,2}$ are fuzzy terms "*about*", "*not about*" associated with the feature of ith sample w_{ij} respectively, then the above linguist fuzzy rules can be written in the following form:

Rule $\mathcal{R}_1$: If t is $m_{1,1}m_{2,1}$, then t is "*iris-setosa*";
Rule $\mathcal{R}_2$: If t is $m_{3,2}m_{2,1}m_{4,2}$, then t is "*iris-setosa*";
Rule $\mathcal{R}$: If t is "$m_{1,1}m_{2,1}$ or $m_{3,2}m_{2,1}m_{4,2}$", then t is "*iris-setosa*".

$\sum_{u=1}^{r} (\prod_{m \in A_u} m)$, which is a formal sum of the sets $\prod_{m \in A_u} m$, $A_u \subseteq M$, is the disjunction of the conjunctions represented by $\prod_{m \in A_u} m$, $u = 1, \ldots, r$. For example, let $A_1 = \{m_{1,1}, m_{2,1}\}$, $A_2 = \{m_{3,2}, m_{2,1}, m_{4,2}\} \subseteq M$, then a new fuzzy set as the disjunction of $\prod_{m \in A_1} m$ and $\prod_{m \in A_2} m$, i.e., "$m_{1,1}m_{2,1}$ or $m_{3,2}m_{2,1}m_{4,2}$", can be represented as

$$\sum_{u=1}^{2} \left(\prod_{m \in A_u} m \right) = \prod_{m \in A_1} m + \prod_{m \in A_2} m.$$

Thus, the fuzzy rule $\mathcal{R}$ can be denoted as follows:

Rule $\mathcal{R}$: If t is $\sum_{u=1}^{2}(\prod_{m\in A_u} m)$, then t is "*iris-setosa*".

The above expressions in **Rule** $\mathcal{R}$ can be formulated as an algebra systems as follows: Let M be a non-empty set. The set EM^* is defined by

$$EM^* = \left\{ \sum_{i\in I}\left(\prod_{m\in A_i} m\right) \mid A_i \subseteq M, i \in I, I \text{ is any no empty indexing set} \right\} \quad (1)$$

An equivalence relation R in EM^* is defined as following:
For $\alpha = \sum_{i\in I}(\prod_{m\in A_i} m), \beta = \sum_{j\in J}(\prod_{m\in B_j} m) \in EM^*$,

$$\alpha R\beta \iff (i)\forall A_i \ (i \in I), \exists B_h(h \in J)$$

such that $A_i \supseteq B_h$; $(ii) \forall B_j \ (j \in J), \exists A_k \ (k \in I)$, such that $B_j \supseteq A_k$.

It's clear that R is an equivalence relation. The quotient set, EM^*/R is denoted by EM. The notation $\sum_{i\in I}(\prod_{m\in A_i} m) = \sum_{j\in J}(\prod_{m\in B_j} m)$ means that $\sum_{i\in I}(\prod_{m\in A_i} m)$ and $\sum_{j\in J}(\prod_{m\in B_j} m)$ are equivalent under equivalence relation R. Thus the semantics they represent are equivalent.

In [16], authors proved that $(EM, \vee, \wedge)$ is completely distributive lattices if the lattice operators $\vee, \wedge$ are defined as following: for any fuzzy sets $\sum_{i\in I}(\prod_{m\in A_i} m)$, $\sum_{j\in J}(\prod_{m\in B_j} m) \in EM$,

$$\sum_{i\in I}\left(\prod_{m\in A_i} m\right) \vee \sum_{j\in J}\left(\prod_{m\in B_j} m\right) = \sum_{k\in I\sqcup J}\left(\prod_{m\in C_k} m\right), \quad (2)$$

$$\sum_{i\in I}\left(\prod_{m\in A_i} m\right) \wedge \sum_{j\in J}\left(\prod_{m\in B_j} m\right) = \sum_{i\in I, j\in J}\left(\prod_{m\in A_i\cup B_j} m\right), \quad (3)$$

where for any $k \in I \sqcup J$ (the disjoint union of I and J, i.e., every element in I and every element in J are always regarded as different elements in $I \sqcup J$), $C_k = A_k$ if $k \in I$, and $C_k = B_k$ if $k \in J$.

In [15], the membership function of $\xi = \sum_{i\in I}\prod_{m\in A_i} m$ is defined as : for any $x \in X$ with weight $\rho_\gamma(u), \gamma \in M, A_i^\tau(x)$ is the set of samples whose degree belonging to $\Pi_{m\in A} m$ is less than that of x,

$$\mu_\xi(x) = \sup_{i\in I} \prod_{\gamma\in A_i} \frac{\sum_{u\in A_i^\tau(x)} \rho_\gamma(u)}{\sum_{u\in X} \rho_\gamma(u)} \quad (4)$$

3 Mining Fuzzy Association Rule based on AFS Fuzzy Logic

In this section, we apply the AFS fuzzy logic to extract the fuzzy rules for the fuzzy reasoning, information process, data analysis and decision systems. In Example 1, we
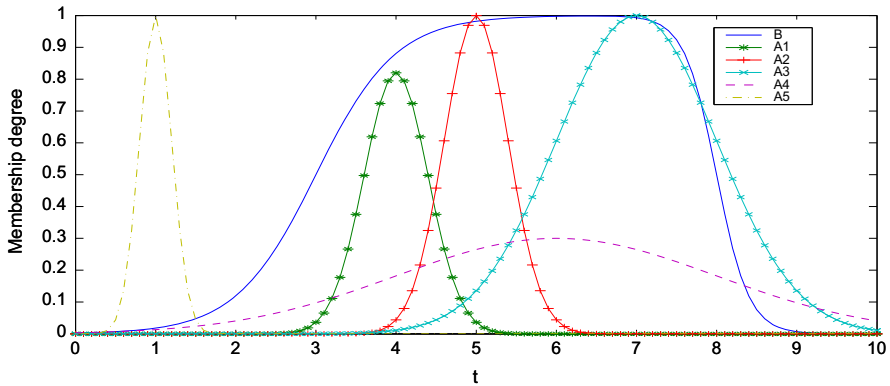
Fig. 1 Membership functions of B, A_i, $i = 1, 2, \ldots, 5$

analyze the degree of implication [7,8], by which the association rules are extended to current fuzzy association rules in the previous approaches to fuzzy association rules.

Example 1 In this example, we study the degree of implication $Dimp$. In the following Fig. 1, $\frac{1}{(1+e^{-2(x-3)})(1+e^{5(x-8)})}$ is the membership function of fuzzy set B, $0.82e^{\frac{-(x-4)^2}{2\times0.4^2}}$ is the membership function of fuzzy set A_1, $e^{\frac{-(x-5)^2}{2\times0.4^2}}$ is the membership function of fuzzy set A_2 , $e^{\frac{-(x-7)^2}{2\times1^2}}$ is the membership function of fuzzy set A_3, $0.3e^{\frac{-(x-6)^2}{2\times2^2}}$ is the membership function of fuzzy set A_4, $e^{\frac{-(x-1)^2}{2\times0.2^2}}$ is the membership function of fuzzy set A_5. $D = \{t_0, t_1, \ldots, t_{100}\}, t_i[I] = 0.1i, i \in [0, 100]$. Let the fuzzy implication operator $FIO(a, b) = 1 - a + ab, a, b \in [0, 1]$, which is one of the operators listed in [7], then by (2) we get the degree of implication as followings: $Dimp(A_1 \Rightarrow B) = 0.9883, Dimp(A_2 \Rightarrow B) = 0.9976, Dimp(A_3 \Rightarrow B) = 0.9567$, $Dimp(A_4 \Rightarrow B) = 0.9664, Dimp(A_5 \Rightarrow B) = 0.9513$.

By Fig. 1, we know the rule: If t is A_5 then t is B is almost false, while rule: If t is A_3 then t is B is almost true, but the the degrees of implication of $Dimp(A_3 \Rightarrow B) = 0.9567, Dimp(A_5 \Rightarrow B) = 0.9513$. It strongly conflicts with our intuition. In the following, we give the new definition of the degree of implication to improve the above drawback of $Dimp$ in [7,8].

Let (I_f, τ, D_f) be the AFS structure for the fuzzy database described in Sect. 2. For fuzzy sets $\alpha, \beta \in EM, \varepsilon \geq 0, \gamma \geq 0$, the degree of implication for $\alpha \Rightarrow \beta$ i.e. "If t is α, then t is β " is defined as following

$$Dimp(\alpha \Rightarrow \beta) = \frac{1}{(1 + \gamma\delta_{\alpha\beta}d_{\alpha\beta})}S(\alpha, \beta), \qquad (5)$$

where $h_\alpha = \max_{t \in D_f}\{\mu_\alpha(t)\}$, $D_\varepsilon = \{t | t \in D_f, \mu_\alpha(t) \geq h_\alpha - \varepsilon\}$, $\delta_{\alpha\beta} = \frac{1}{|D_\varepsilon|}\sum_{t \in D_\varepsilon}|\mu_\beta(t) - \mu_\alpha(t)|$, $d_{\alpha\beta} = \frac{1}{|D_\varepsilon|}\sum_{t \in D_\varepsilon}|h_\beta - \mu_\beta(t)|$, $S(\alpha, \beta)$ is the ratio

subsethood [18] defined in the framework of AFS as following:

$$S(\alpha, \beta) = \frac{c(\alpha \wedge \beta)}{c(\alpha)} = \frac{\sum_{t \in D_f} \mu_{\alpha \wedge \beta}(t)}{\sum_{t \in D_f} \mu_\alpha(t)}. \tag{6}$$

In order to compare the definition of $Dimp$ (5) with $Dimp$ in [7,8], let $\mu_{\alpha \wedge \beta}(t) = \min\{\mu_\alpha(t), \mu_\beta(t)\}$ in (6), but we should notice that in AFS theory $\mu_{\alpha \wedge \beta}(t)$ is determined by the distribution of the original data, instead of t-norm min. Let $\varepsilon = 0.5$, $\gamma = 5$, by (5), we get $Dimp$ for the fuzzy sets in Example 1 as followings: $Dimp(A_1 \Rightarrow B) = 0.8603$, $Dimp(A_2 \Rightarrow B) = 0.9835$, $Dimp(A_3 \Rightarrow B) = 0.8469$, $Dimp(A_4 \Rightarrow B) = 0.4750$, $Dimp(A_5 \Rightarrow B) = 0.0106$. By Fig. 1, we can observe that $Dimp'$s obtained by (5) are very similar to our intuition.

It is obvious that we can apply (5) to find the degree of implication $Dimp(\alpha \Rightarrow \beta)$ for any pair fuzzy concepts $\alpha, \beta \in E(I_f)$ and $Dimp(\alpha \Rightarrow \beta)$ in (5) is determined by the distribution of the database, instead of t-norm, t-conorm, fuzzy implication operator. The following Example 2 will show that the proposed degree of implication in the framework of AFS theory is an objective reflection of the fuzzy logic relations among the fuzzy concepts in the database and is also similar to our intuition.

In general, $I = \{I_1, I_2, \ldots, I_m\}$ be the set of items where each I_j $(1 \leq j \leq m)$ is an attribute of the original quantitative database $D = \{t_1, t_2, \ldots, t_n\}$, and q_j be the number of fuzzy sets defined on the domain of I_j. Then the original quantitative database D can be transformed into an extended fuzzy database $D_f = \{t_1, t_2, \ldots, t_n\}$ with the set of items $I_f = \{m_1, m_2, \ldots, m_q\}$, where m_i is the fuzzy concept like that in Example 1. Let $X, Y \subset I$, $X \cap Y = \varnothing$, the elements of X are conditional attributes and the elements of Y are decision attributes. Let M_X be the set of the fuzzy concepts in I_f associated with the attributes in X and M_Y be the set of the fuzzy concepts in I_f associated with the attributes in Y. In the real world applications, we need to find the fuzzy rules: If t is α, then t is β, $\alpha \in E(M_X)$, $\beta \in E(M_Y)$. In theory, we can apply (5) to check $Dimp(\alpha \Rightarrow \beta)$ for each pair of α, β, but in practice, since there are more than $\sum_{i=1}^{|M_X|}(2^{C_n^i} - 1)$ elements in $E(M_X)$, hence it is impossible for us to check each one. Fortunately, this problem most likely be solved by applications of the perfect mathematical property of lattice $(E(M_X), \vee, \wedge, ')$ in future. In the following, we just propose an approximate algorithm of mining fuzzy rules: "If t is α, then t is $\beta, \alpha \in E(M_X)$" for a given fuzzy concept β in $E(M_Y)$:

The fuzzy association rule mining algorithm for a given fuzzy concept on decision attributes

1 ζ: $\zeta = \{l\}, l \in M_X$, $Dimp(\{l\} \Rightarrow \beta) = \max_{m \in M_X}\{Dimp(\{m\} \Rightarrow \beta)\}$

2 s: $s \in D_f$, $v = \mu_\zeta(s) - \mu_\beta(s) = \max_{t \in D_f}\{\mu_\zeta(t) - \mu_\beta(t)\}$

3   **for** $\{v \geq \delta_1\}$ **do** ($\delta_1 > 0$ i.e. $\max_{t \in D_f}\{\mu_\alpha(t) - \mu_\beta(t)\} < \delta_1$ for the final rule: "If t is α, then t is β")

4       $M_s = \{m|m \in M_X, \mu_{\{m\}}(s) < \mu_\beta(s) + \delta_1\}$

5       $\zeta = \zeta \wedge \xi$: $\xi \in M_s$, $Dimp(\zeta \wedge \xi \Rightarrow \beta) = \max_{\eta \in M_s}\{Dimp(\zeta \wedge \eta \Rightarrow \beta)\}$

6   s: $s \in D_f$, $v = \mu_\zeta(s) - \mu_\beta(s) = \max_{t \in D_f}\{\mu_\zeta(t) - \mu_\beta(t)\}$

7   **endfor**

8 k: $k \in D_f$, $u = \mu_\zeta(k) - \mu_\beta(k) = \max_{t \in D_f}\{\mu_\beta(t) - \mu_\zeta(t)\}$

9   **for** $\{u \geq \delta_2\}$ **do** ($\delta_2 > 0$ i.e. $\max_{t \in D_f}\{\mu_\beta(t) - \mu_\alpha(t)\} < \delta_2$ for the final rule: "If t is α, then t is β")

10    $M_k = \{m | m \in M_X, \mu_{\{m\}}(k) > \mu_\beta(k) + \delta_2\}$

11    $\xi: \xi \in M_k, Dimp(\xi \Rightarrow \beta) = \max_{\eta \in M_k}\{Dimp(\eta \Rightarrow \beta)\}$

12    $s: s \in D_f, v = \mu_\xi(s) - \mu_\beta(s) = \max_{t \in D_f}\{\mu_\xi(t) - \mu_\beta(t)\}.$

13   **for** $\{v \geq \delta_1\}$ **do**

14        $M_s = \{m | m \in M_X, \mu_{\{m\}}(s) < \mu_\beta(s) + \delta_1\}$

15        $\xi = \xi \wedge \varsigma: \varsigma \in M_k, Dimp(\zeta \wedge \varsigma \Rightarrow \beta) = \max_{\eta \in M_s}\{Dimp(\zeta \wedge \eta \Rightarrow \beta)\}$

16        $s: s \in D_f, v = \mu_\xi(s) - \mu_\beta(s) = \max_{t \in D_f}\{\mu_\xi(t) - \mu_\beta(t)\}$

17   **endfor**

18    $\zeta = \zeta \vee \xi$

19    $k: k \in D_f, u = \mu_\zeta(k) - \mu_\beta(k) = \max_{t \in D_f}\{\mu_\beta(t) - \mu_\zeta(t)\}$

20   **endfor**

21   "If t is ζ, then t is β" is a rule for the given fuzzy sets $\beta \in E(M_Y)$ such that $Dimp(\zeta \Rightarrow \beta)$ is an estimate of $\max_{\eta \in E(M_X)}\{Dimp(\eta \Rightarrow \beta)\}$ and $\forall t \in D_f$, $-\delta_1 < \mu_\beta(t) - \mu_\zeta(t) < \delta_2$.

In the following example we apply the proposed algorithm to the Iris dataset.

Example 2 For the Iris dataset, let (I_f, τ, D_f) be the AFS structure, (I_f, τ, D_f, S) be the semi-cognitive field. $\forall m \in I_f, \rho_m: D_f \rightarrow [0, \infty)$. For any concept $\alpha \in E(I_f)$, its membership function is defined in [19]. We apply the above algorithm to find rules for class iris-setosa, iris-versicolor, iris-virginica i.e. $M_X = \{m_1, m_2, \ldots, m_{24}\}$, $M_Y = \{m_{25}, m_{26}, \ldots, m_{30}\}$,

if t is α, then t is $\{m_{25}\}, \alpha \in E(M_X)$, $\{m_{25}\}$ is the crisp concept: class iris-setosa;

if t is α, then t is $\{m_{27}\}, \alpha \in E(M_X)$, $\{m_{27}\}$ is the crisp concept: class iris-versicolor;

if t is α, then t is $\{m_{29}\}, \alpha \in E(M_X)$, $\{m_{29}\}$ is the crisp concept: iris-virginica.

Let $\varepsilon = 0.5, \gamma = 10$, in (5) and $\delta_1 = 0.01, \delta_2 = 0.5$ in the algorithm for finding rule for class iris-setosa, $\delta_1 = 0.2, \delta_2 = 0.5$ in the algorithm for finding rules for classes iris-versicolor, iris-virginica. We get the following fuzzy rules:

class iris-setosa: If t is ζ_1 then t is $\{m_{25}\}$, where $\zeta_1 = \{m_{13}\} + \{m_{19}\} + \{m_4, m_9, m_{22}, m_6, m_{11}, m_{24}\}$.

class iris-versicolor: If t is ζ_2 then t is $\{m_{27}\}$, where $\zeta_2 = \{m_1, m_9, m_{21}\} + \{m_8, m_3, m_{15}, m_{21}, m_5\} + \{m_8, m_3, m_9, m_{15}, m_{21}\} + \{m_3, m_9, m_{15}, m_{21}, m_{11}\} + \{m_2, m_{15}, m_{21}, m_5, m_{11}\}$.

iris-virginica: If t is ζ_3 then t is $\{m_{29}\}$, where $\zeta_3 = \{m_4, m_{12}, m_{23}\} + \{m_{14}, m_{20}, m_3, m_9, m_{11}, m_{17}, m_{23}\} + \{m_2, m_{14}, m_{20}, m_4, m_9, m_{11}\} + \{m_7, m_{14}, m_{20}, m_3, m_5, m_{17}\} + \{m_8, m_{14}, m_{20}, m_3, m_5, m_{23}\} + \{m_8, m_{14}, m_3, m_9, m_{21}, m_5, m_{17}\} + \{m_{14}, m_{20}, m_3, m_9, m_5, m_{11}, m_{23}\} + \{m_2, m_{14}, m_{20}, m_9, m_5, m_{11}, m_{17}, m_{23}\} + \{m_{22}, m_{17}\}$.

The following Fig. 2 shows the membership functions for fuzzy concepts $\zeta_1, \zeta_2, \zeta_3 \in E(M_X)$ as the antecedents for the fuzzy rules of class 1, 2, 3 (i.e. $k = 1$ iris-setosa, $k = 2$ iris-versicolor or $k = 3$ virginica), respectively.

For any $t \in D_f$, t is class k, if $\mu_{\zeta_k}(t) = \max\{\mu_{\zeta_1}(t), \mu_{\zeta_2}(t), \mu_{\zeta_3}(t)\}$. For the total 150 samples, there are 3 classifying error samples (i.e. reclassification accuracy 98 %) and they are list in Table 1.

In [11], the authors have used the Iris example for pedagogical reasons to compare the current rule-extraction systems and the reclassification accuracies (in-sample
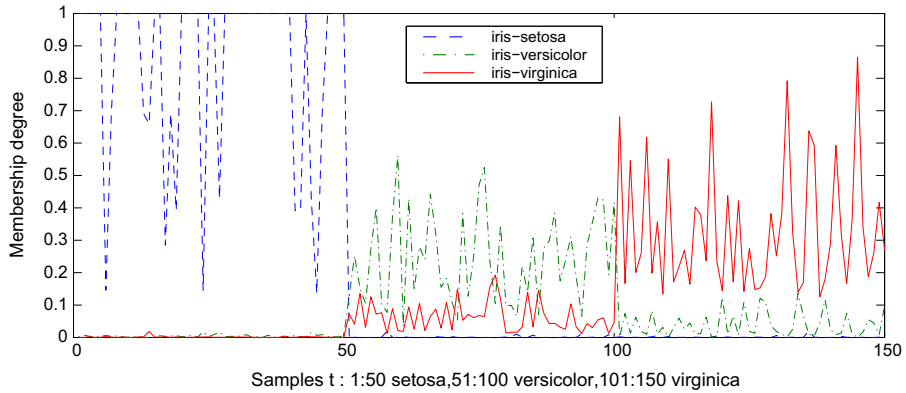
Fig. 2 The membership functions of ζ_1, ζ_2, ζ_3

Table 1 The error classifying samples by the fuzzy rules

no. sample	71	78	86
$\mu_{\zeta_1}(.)$	0.0000	0.0001	0.0001
$\mu_{\zeta_2}(.)$	0.0541	0.1052	0.0705
$\mu_{\zeta_3}(.)$	0.1497	0.1925	0.1464

Table 2 Number of rules(NR), number of conditions (NC), and Number of features used (NF) by rules extracted for the Iris dataset by different systems

Method	NR/NC/NF	Type	Reclass. accuracy (%)
ReFuNN[20]	9/26/4	F	95.7
ReFuNN[20]	104/36/4	F	95.7
NEFCLASS[21]	7/28/4	F	96.7
NEFCLASS[21]	3/6/2	F	96.7
FuNe-I[22]	7/-/3	F	96.0
AFS	3/16/4	F	98.0
Grobian[23]	118/-/4	R	100.0
GA+NN[24]	6/6/4	W	100.0
C-MLP2LN[25]	2/2/1	C	95.7
C-MLP2LN[25]	2/2/2	C	96.0
C-MLP2LN[25]	2/3/2	C	98.0
SSV[25]	2/2/2	C	98.0

Rules Are Either Crisp (C), Fuzzy (F), Rough (R), or Weighted (W)

accuracies for the whole dataset) of rules derived by several rule-extraction systems were reported in Table 2. We also add the results in Example 2 named as AFS in Table 2.

By Table 2, one knows that the proposed fuzzy association rule mining algorithm obtains the best result of all fuzzy algorithm and the degrees of implication *Dimp* defined by (5) in the framework of AFS theory is not only over come the obvious

drawback of the other definitions of degree of support $DSupp$, implication $Dimp$ and confidence for fuzzy mining association rules, which are based on t-norm, fuzzy implication operator and fuzzy similarity relation given in advance, but also provides a promising fuzzy mining approach for real world applications.

4 Conclusion

By degrees of implication and the mining algorithm proposed, we know that the AFS fuzzy logic is more appropriate for human thinking and natural language than other fuzzy theory and the fuzzy rules automatically and directly extracted from the database are simple, accuracy and interpretable. Since each extracted fuzzy rule by the proposed algorithm has definite semantic signification expressed by the simple fuzzy sets on the attributes which are determined by the special distribution of each data systems, instead of human intuition, hence the rules are objective representations of the knowledge hidden in the data systems. AFS fuzzy logic is an important mathematical tools for us to study law of human thinking and to apply computers to do more intelligent works for human. We hope that more mathematicians, scientists and engineers can pay attention to AFS theory and the AFS fuzzy logic.

References

1. Luger GF (2005) Artificial intelligence, 5th edn. Addison Wesley, Reading
2. Anderson TW (1984) An introduction to multivariate statistical analysis, 2nd edn. Wiley, New York
3. Arnold SF (1990) Mathematical statistics. Prentice Hall, Englewood Cliffs, NJ
4. Agrawal R, Imielinski T, Swami A (1993) Mining association rules between sets of items in large databases, In: Proceeding of the ACM SIGMOD Conference on Management of Data, Washington, DC, pp 207–216
5. Han J, Kamber M (2000) Data mining: concepts and techniques. Morgan Kaufmann Publishers, Amsterdam
6. Srikant R, Agrawal R (1996) Mining quantitative association rules in large relational tables. In:SIGMOD 96 6/96 Montreal
7. Yan P, Chen GQ (2005) Discovering a cover set of ARsi with hierarchy from quantitative databases. Inf Sci 173:319C336
8. Chen GQ, Wei Q (2002) Fuzzy association rules and the extended mining algorithms. Inf Sci 147:201–228
9. Delgado M, Marin N, Sachez D, Vila MA (2003) Fuzzy association rules: general model and applications. IEEE Trans Fuzzy Syst 11(2):214–225
10. Berzal F, Blanco I, Sánchez D, Serrano JM, Vila MA (2005) A definition for fuzzy approximate dependencies. Fuzzy Sets Syst 149:105–129
11. Duch W, Setiono R, Zurada JM (2004) Computational intelligence methods for rule-based data understanding. Proc IEEE 92(5):711–805
12. Au WH, Chan KCC (2003) Mining fuzzy association rules in a bank-account database. IEEE Trans Fuzzy Syst 11(2):238–248
13. Shu YJ, Tsang E, Yeung DS, Shi D (2000) Mining fuzzy association rules with weighted items. In: Proceedings of the IEEE international conference on system, man and cybernetics (SMC2000), Nashville, Tennessee
14. Hullermeier E (2001) Implication-based fuzzy association rules. ECML/PKDD, Freiburg
15. Liu XD, Pedrycz W (2009) AFS theory and its applications. Springer-Verlag, Heidelberg
16. Liu XD (1998) The fuzzy theory based on AFS algebras and AFS structure. J Math Anal Appl USA 217:459–478

17. Mertz J, Murphy PM (1996) UCI repository of machine learning databases. Available: http://www.ics.uci.edu/pub/machinelearning-data-bases
18. Kosko B (1997) Fuzzy engineering. Prentice Hall, Upper Saddle River
19. Liu XD, Wang W, Chai TY (2005) The fuzzy clustering analysis based on AFS theory. IEEE Trans Syst Man Cybern Part B 35(5):1013–1027
20. Kasabov N (1996) Foundations of neural networks, fuzzy systems and knowledge engineering. MIT Press, Cambridge, MA
21. Nauck D, Nauck U, Kruse R (1996) Generating classification rules with the neuro-fuzzy system NEFCLASS. In: Proceedings of the Biennial conference of the North American Fuzzy Information Processing Society NAFIPS. Berkeley, CA. vol. 96, pp 466–470
22. Halgamuge SK, Glesner M (1994) Neural networks in designing fuzzy systems for real world applications. Fuzzy Sets Syst 65:1–12
23. Browne C, Duntsch I, Gediga G (1998) IRIS revisited: a comparison of discriminant and enhanced rough set data analysis. In: Polkowski L, Skowron A (eds) Rough sets in knowledge discovery, vol 2. Physica-Verlag, Heidelberg, pp 345–368
24. Jagielska I, Matthews C, Whitfort T (1996) The application of neural networks, fuzzy logic, genetic algorithms and rough sets to automated knowledge acquisition, In: Proceeding of the 4th international conference soft computing, vol. 2, pp 565–569
25. Duch W, Adamczak R, Grabczewski K (2001) A new methodology of extraction, optimization and application of crisp and fuzzy logical rules. IEEE Trans Neural Netw 12:277–306