NEW APPROACHES TO THE FUZZY CLUSTERING VIA AFS THEORY

YAN REN, MINGLI SONG, AND XIAODONG LIU

Abstract. In this paper, we study and improve the fuzzy clustering index and clustering algorithm proposed by X. Liu et. al. in IEEE Transactions on Systems, Man, Cybernetics, 2005, and also study the fuzzy clustering problems by the order relation descriptions of the objects on each attribute, instead of the real number descriptions. The new fuzzy clustering algorithm has two main advantages: One is that it can mimic the human reasoning processes and offer an interpretable clustering algorithm which is represented by some fuzzy sets with definite semantic interpretations. Another is the data types of the attributes can be various types or sub-preference relations, even descriptions of human intuitions. Finally, the well known real-world iris data set is used to illustrate accuracy of the new clustering algorithm. A number of illustrative examples show that this approach offers a far more flexible and effective means for the intelligent systems in real-world applications.

Key Words. AFS algebras, Molecular lattices, Fuzzy matrices, AFS structures, Sub-preference relations.

1. Introduction

Today, it is an era for mass data storage. However, knowledge representation and acquisition for huge data is a bottleneck for engineering applications. Currently many researchers are working hard to extract useful information (i.e., pattern recognition) from data and then classify the information. Due to endless efforts, there are many approach aimed alleviating this problem including fuzzy set theory [11]. The incorporation of fuzzy sets into the representations of fuzzy concepts enables us to combine the uncertainty handling and approximate reasoning capabilities with the comprehensibility and ease of application.

In many fuzzy theories, the membership functions are often given by personal intuition and the logic operations are implemented by a kind of triangular norms, or shortly t-norm which is chosen in advance and independent of the distribution of the raw data. The large-scale intelligence systems in real-world applications are usually very large and complex, containing such a large number of concepts that it is impossible or difficult to define the membership functions just by personal intuition and to choose suitable one from infinite kinds of triangular norms to implement fuzzy logic operations.

In [18-24], the author has proposed AFS (Axiomatic Fuzzy Set) theory in which fuzzy sets (membership functions) and their logic operations are impersonally and automatically determined by a consistent algorithm according to the distributions of original data. In [25], the complement operation in AFS theory has been obtained,
thus a fuzzy logic system, which is called AFS fuzzy logic, has been developed. Recently, AFS theory has been developed further and applied to fuzzy decision trees [26], fuzzy clustering analysis [27], concept representations [28], fuzzy cognitive maps [29] and credit rating analysis [30]. In AFS theory, the membership functions and their logic operations are determined by AFS structures and AFS algebras. An AFS structure is a mathematical object of triple $(M, \tau, X)$ which is a special family of combinatorics objects [7], where $X$ is the set of objects, $M$ is a set of some simple concepts [27] on $X$ and $\tau$ is mathematical abstract of the complex relations of objects concerning the original data and facts such as database, sub-preference relations [27], even human intuition descriptions. AFS algebra is a family of molecular lattices (i.e. completely distributive lattices) generated by sets such as $X, M$. With the AFS structure $(M, \tau, X)$, which can be directly obtained from the given dataset, any concepts on $X$ can be represented by the fuzzy sets in $EM$, which is the AFS algebra over $M$. Using AFS algebra and AFS structure $(M, \tau, X)$, a great large number of complex fuzzy concepts on $X$ and their logic operations can be implemented by the algebraic operators on few simple concepts in $M$. The membership functions and their logic operations are more accurate and impersonal reflections of the original data and facts than given by human intuition in other fuzzy theories, and the information contained in the original data and facts is preserved at significant extent (please refer to the survey in [27,31]).

Clustering of objects are important areas in a variety of fields, such as pattern recognition, artificial intelligence and vision analysis. In this paper, we mainly focus on the fuzzy clustering. Conventional non-fuzzy or crisp classification techniques assume that a pattern $\alpha$ belongs to only one class. In real world applications, a pattern $\alpha$ always belongs to more than one class at different degrees. It is natural to apply fuzzy set theory in cluster analysis. Bellman et al. [14] and Ruspini [2] first initiated research in fuzzy clustering. Now fuzzy clustering has been widely studied and applied in diverse areas. See, for example, Keller [5], Bezdek [6,8,9], Pedrycz [16,17], Yang et al. [12,13]. In [27], the authors have proposed a new fuzzy clustering algorithm in the framework of AFS theory. Compared with the current fuzzy clustering algorithm, the new algorithm has the following advantages: 1. The attributes of objects in it can be various data types or sub-preference relations, even human intuition descriptions. 2. The distance function and objective function are not required, and the cluster number or the class label need not be given beforehand. 3. Each cluster is described by a fuzzy set in $EM$, which is the AFS fuzzy logical compound of the simple attributes on some features with definite semantic interpretation and determines the degree of each pattern belonging to this cluster. In this paper, an exhaustive study of the fuzzy clustering algorithm and the fuzzy clustering index proposed in [27] is carried out by the applications of the real world Iris dataset (150 samples), which is larger than the illustrative examples (8, 10 samples) in [27], and some improvements of the fuzzy clustering index and algorithm have been made.

This paper is organized as follows: Section 2 introduces some basic concepts and presents several pertinent results on AFS theory. Section 3 improves the fuzzy clustering algorithm and the fuzzy clustering index proposed in [27]. Section 4 applies this algorithm to a real-world iris data set, and the experimental results demonstrate the effectiveness of the proposed clustering algorithm. Section 5 concludes this paper.
2. Preliminaries of the AFS theory

In this section, we recall the notations and definitions, present several pertinent results of AFS theory concerning the fuzzy clustering analysis and classification. We employ the notations, definitions and the symbols in [27] in what follows. In essence, the AFS framework supports the studies on how to convert the information in the training examples or databases into the membership functions and their fuzzy logic operations. AFS theory is made of AFS structures which is a special kind of combinatorics objects and AFS algebra which is a family of completely distributivity lattices. About the detail mathematical properties of AFS structure and AFS algebra, please see [18, 22, 27, 31].

In general, we explain the fuzzy sets and crisp subsets on $X$ as the followings:

For a fuzzy set $\zeta$ on universe of discourse $X$, any $x \in X$, either $x$ belongs to $\zeta$ at some degree or does not belong to $\zeta$ at all. While for a crisp subset $A$ of $X$, any $x \in X$, either $x$ belongs to $A$ or does not belong to $A$ at all.

Based on this statement, both a fuzzy set and a crisp subset on $X$ can be represented by a binary relation $R$ on $X$ through comparing the degrees of each pair of $x, y$ in $X$ belonging to the concept as described below.

**Definition 1.** ([27]) Let $\zeta$ be any concept on the universe of discourse $X$. $R_\zeta$ is called a binary relation (i.e., $R_\zeta \subset X \times X$) of $\zeta$ if $R_\zeta$ satisfies:

1. $x, y \in X, (x, y) \in R_\zeta \iff x$ belongs to concept $\zeta$ at some degree and the degree of $x$ belonging to $\zeta$ is larger than or equal to that of $y$, or $x$ belongs to concept $\zeta$ at some degree and $y$ does not at all.

Although fuzzy concepts are ambiguous, the binary relations corresponding to them are crisp subset of $X \times X$. In real world applications, $R_\zeta$ can also be obtained by comparing the degrees of each pair objects belonging to concept $\zeta$ through human intuitions without necessary to represent them in $[0, 1]$ or a lattice in advance. We should notice that $(x, x) \in R_\zeta$ means that $x$ belongs to concept $\zeta$ at some degree and $(x, x) \notin R_\zeta$ means that $x$ does not belong to concept $\zeta$ at all. For example, let $X$ be a set of persons and $m$ be the simple concept “tall”, even if we do not know the exact height of $x$ and $y$, the degrees of $x$ and $y$ belonging to a simple concept $m$ can be compared. Now that each human concept corresponds a unique binary relation, how can we study human concepts by binary relations? First we should dealt with a class of simple concepts whose membership functions and logic operations are simple enough to be obtained, then the complex concepts are represented by these simple concepts.

**Definition 2.** ([27]) Let $X$ be a set and $R$ be a binary relation on $X$. $R$ is called a sub-preference relation on $X$ if for $x, y, z \in X$, $x \neq y$, $R$ satisfies the following conditions:

1. $D2-1. \text{If} (x, y) \in R, \text{then} (x, x) \in R$;
2. $D2-2. \text{If} (x, x) \in R \text{and} (y, y) \notin R, \text{then} (x, y) \in R$;
3. $D2-3. \text{If} (x, x), (y, z) \in R, \text{then} (x, z) \in R$;
4. $D2-4. \text{If} (x, x) \in R \text{and} (y, y) \in R, \text{then either} (x, y) \in R \text{or} (y, x) \in R$.

A concept $\zeta$ is called a simple concept or simple attribute on $X$ if $R_\zeta$ is a sub-preference relation. Otherwise $\zeta$ is called a complex concept or a complex attribute on $X$. 
Let's give an example of complex concepts. Let $X$ be a set of persons and cars. $x, y \in X$, $x$ is a person and $y$ is a car. If we consider concept “beautifull”, then the degrees of $x, y$ belonging to “beautifull” can not be compared although both $x$ and $y$ belong to “beautifull” at some degree, i.e., $(x, x), (y, y) \in R_{beautifull}$, $(x, y) \notin R_{beautifull}$. This implies that $4$ of Definition $2$ is not satisfied and “beautifull” is a complex concept on $X$. Whether a concept is a simple concept or a complex concept is determined by both the universe of discourse and the intent of a concept. By any simple concept $\zeta$, $X$ is divided into three disjoint parts as follows:

$$
T_{\zeta} = \{x \in X | (x, y) \in R_{\zeta}, \forall y \in X\}, \\
F_{\zeta} = \{x \in X | (x, x) \notin R_{\zeta}\}, \\
M_{\zeta} = X - T_{\zeta} - F_{\zeta}.
$$

The degree of each element in $T_{\zeta}$ or $F_{\zeta}$ belonging to concept $\zeta$ is $1$ or $0$ respectively if concept $\zeta$ is represented by an ordinary fuzzy set and is maximum element of lattice $L$ or minimum element respectively if concept $\zeta$ is represented by a L-fuzzy set. In other words, absolutely true or false. The elements in $M_{\zeta}$ belong to concept $\zeta$ at different degrees in open interval $(0, 1)$ if concept $\zeta$ is represented by an ordinary fuzzy set. The elements in $M_{\zeta}$ form a linear ordered chain in lattice $L$ if concept $\zeta$ is represented by a L-fuzzy set (4 of Definition 2). Concept $\zeta$ is a crisp concept if $M_{\zeta} = \emptyset$.

In [18], the author has defined a family of molecular lattices, the AFS algebras, i.e., the $EI, EII, ..., EI^n$ algebras and applied AFS algebras to study the lattice valued representations for fuzzy concepts.

**Definition 3.** ([18]) Let $X_1, ..., X_n, M$ be $n + 1$ non-empty sets. Then the set $EX_1...X_nM^*$ is defined by

$$EX_1...X_nM^* = \{\sum_{i \in I}(u_{1i}...u_{ni}A_i) | A_i \in 2^M, u_{ri} \in 2^{X_r}, r = 1, ..., n, i \in I, \\
I \text{ is a non-empty indexing set}\}.
$$

In the case $n = 0,$

$$EM^* = \{\sum_{i \in I} A_i | A_i \in 2^M, i \in I, I \text{ is a non-empty indexing set}\},
$$

where the element $\sum_{i \in I}(u_{1i}...u_{ni}A_i)$ is composed of terms $(u_{1i}...u_{ni}A_i)$’s, $i \in I$, separated by “+”. $\sum_{i \in I}(u_{1i}...u_{ni}A_i)$ and $\sum_{i \in I}(u_{1i}...u_{ni}A_i)$ are the same element of $EX_1...X_nM^*$ if $p$ is a bijection from $I$ to $I$. When $I$ is finite, $\sum_{i=1}^n(u_{1i}...u_{ni}A_i)$ is also denoted as $(u_{11}...u_{ni}A_i) + ... + (u_{1g}...u_{ng}A_i)$.

For a set, we know that the subsets of a set often contain or represent some useful information and knowledge. In real world applications, instead of one set, often many sets are involved and the information and knowledge represented by the subsets of different sets may have some kinds of relations. In order to study the complicated relations among the information and knowledge associated to different sets, we introduce the notation of $EX_1...X_nM^*$. Every element of $EX_1...X_nM^*$ is a “formal sum” of the terms constituted by the subsets of $X_1$, $X_2$, ..., $X_n$, $M$. For $\gamma = \sum_{i \in I}(u_{1i}...u_{ni}A_i) \in EX_1...X_nM^*$, $\gamma$ can be regarded as the “synthesis” of the information represented by all terms $u_{1i}...u_{ni}A_i$’s. In practice, $M$ is a set of simple concepts, and $X_1$, $X_2$, ..., $X_n$ are the sets associated the concepts in $M$. For example, let $X$ be a set of persons and $M$ be a set of concepts such as “male”, “female”, “age”, “height”, “salary”, “hair black”, “hair white”....etc. For $\sum_{i \in I}(u_iA_i) \in EM^*$, every term $u_iA_i$, $i \in I$, may mean that the persons in set $u_i \subset X$ satisfy some “condition” described by the attributes in $A_i \subset M$. AFS theory supports studies on how to convert the information represented by

\[
\begin{align*}
\text{Let's give an example of complex concepts. Let } X \text{ be a set of persons and cars. } x, y \in X, x \text{ is a person and } y \text{ is a car. If we consider concept “beautifull”, then the degrees of } x, y \text{ belonging to “beautifull” can not be compared although both } x \\
\text{and } y \text{ belong to “beautifull” at some degree, i.e., } (x, x), (y, y) \in R_{beautifull}, (x, y) \notin R_{beautifull}. \text{ This implies that } 4 \text{ of Definition } 2 \text{ is not satisfied and “beautifull” is a complex concept on } X. \text{ Whether a concept is a simple concept or a complex concept is determined by both the universe of discourse and the intent of a concept. By any simple concept } \zeta, X \text{ is divided into three disjoint parts as follows:}
\end{align*}
\]

\[
T_\zeta = \{x \in X | (x, y) \in R_\zeta, \forall y \in X\}, \\
F_\zeta = \{x \in X | (x, x) \notin R_\zeta\}, \\
M_\zeta = X - T_\zeta - F_\zeta.
\]

The degree of each element in $T_\zeta$ or $F_\zeta$ belonging to concept $\zeta$ is 1 or 0 respectively if concept $\zeta$ is represented by an ordinary fuzzy set and is maximum element of lattice $L$ or minimum element respectively if concept $\zeta$ is represented by a L-fuzzy set. In other words, absolutely true or false. The elements in $M_\zeta$ belong to concept $\zeta$ at different degrees in open interval $(0, 1)$ if concept $\zeta$ is represented by an ordinary fuzzy set. The elements in $M_\zeta$ form a linear ordered chain in lattice $L$ if concept $\zeta$ is represented by a L-fuzzy set (4 of Definition 2). Concept $\zeta$ is a crisp concept if $M_\zeta = \emptyset$.

In [18], the author has defined a family of molecular lattices, the AFS algebras, i.e., the $EI, EII, ..., EI^n$ algebras and applied AFS algebras to study the lattice valued representations for fuzzy concepts.

**Definition 3.** ([18]) Let $X_1, ..., X_n, M$ be $n + 1$ non-empty sets. Then the set $EX_1...X_nM^*$ is defined by

$$EX_1...X_nM^* = \{\sum_{i \in I}(u_{1i}...u_{ni}A_i) | A_i \in 2^M, u_{ri} \in 2^{X_r}, r = 1, ..., n, i \in I, \\
I \text{ is a non-empty indexing set}\}.
$$

In the case $n = 0,$

$$EM^* = \{\sum_{i \in I} A_i | A_i \in 2^M, i \in I, I \text{ is a non-empty indexing set}\},
$$

where the element $\sum_{i \in I}(u_{1i}...u_{ni}A_i)$ is composed of terms $(u_{1i}...u_{ni}A_i)$’s, $i \in I$, separated by “+”. $\sum_{i \in I}(u_{1i}...u_{ni}A_i)$ and $\sum_{i \in I}(u_{1i}...u_{ni}A_i)$ are the same element of $EX_1...X_nM^*$ if $p$ is a bijection from $I$ to $I$. When $I$ is finite, $\sum_{i=1}^n(u_{1i}...u_{ni}A_i)$ is also denoted as $(u_{11}...u_{ni}A_i) + ... + (u_{1g}...u_{ng}A_i)$.

For a set, we know that the subsets of a set often contain or represent some useful information and knowledge. In real world applications, instead of one set, often many sets are involved and the information and knowledge represented by the subsets of different sets may have some kinds of relations. In order to study the complicated relations among the information and knowledge associated to different sets, we introduce the notation of $EX_1...X_nM^*$. Every element of $EX_1...X_nM^*$ is a “formal sum” of the terms constituted by the subsets of $X_1$, $X_2$, ..., $X_n$, $M$. For $\gamma = \sum_{i \in I}(u_{1i}...u_{ni}A_i) \in EX_1...X_nM^*$, $\gamma$ can be regarded as the “synthesis” of the information represented by all terms $u_{1i}...u_{ni}A_i$’s. In practice, $M$ is a set of simple concepts, and $X_1$, $X_2$, ..., $X_n$ are the sets associated the concepts in $M$. For example, let $X$ be a set of persons and $M$ be a set of concepts such as “male”, “female”, “age”, “height”, “salary”, “hair black”, “hair white”....etc. For $\sum_{i \in I}(u_iA_i) \in EM^*$, every term $u_iA_i$, $i \in I$, may mean that the persons in set $u_i \subset X$ satisfy some “condition” described by the attributes in $A_i \subset M$. AFS theory supports studies on how to convert the information represented by
the elements of $EX_1...X_n M^*$ for the training examples and databases into the membership functions and their fuzzy logic operations.

**Definition 4.** ([18]) Let $X_1, ..., X_n, M$ be $n+1$ non-empty sets. A binary relation $R$ on $EX_1...X_n M^*$ is defined as follows. For any $\sum_{i \in I}(u_{i1}...u_{in}A_i), \sum_{j \in J}(v_{j1}...v_{jn}B_j) \in EX_1...X_n M^*$,

$$\left[\sum_{i \in I}(u_{i1}...u_{in}A_i) \right] R \left[\sum_{j \in J}(v_{j1}...v_{jn}B_j) \right] \iff$$

(i) $\forall (u_{i1}...u_{in}A_i) (i \in I), \exists (v_{j1}...v_{jn}B_j) (h \in J)$ such that $A_i \supseteq B_h$, $u_{ri} \subseteq v_{rh}$, $1 \leq r \leq n$;

(ii) $\forall (v_{j1}...v_{jn}B_j) (j \in J), \exists (u_{i1}...u_{in}A_i) (k \in I)$, such that $B_j \supseteq A_k$, $v_{rj} \subseteq u_{rk}$, $1 \leq r \leq n$.

It’s obvious that $R$ is an equivalence relation. Thus $EX_1...X_n M^*/R$, the quotient set, is well-defined and denoted by $EX_1...X_n M$. By using notation $\sum_{i \in I}(u_{i1}...u_{in}A_i) = \sum_{j \in J}(v_{j1}...v_{jn}B_j)$, we mean that $\sum_{i \in I}(u_{i1}...u_{in}A_i)$ and $\sum_{j \in J}(v_{j1}...v_{jn}B_j)$ are equivalent under equivalence relation $R$.

**Proposition 1.** ([18]) Let $X_1, ..., X_n, M$ be $n+1$ non-empty sets. If $A_i \subseteq A_i$, $u_{ri} \supseteq u_{rs}$, $r = 1, 2, ..., n$, $t, s \in I, t \neq s$, $\sum_{i \in I}(u_{i1}...u_{in}A_i) \in EX_1...X_n M^*$, then

$$\sum_{i \in I}(u_{i1}...u_{in}A_i) = \sum_{i \in I-s}(u_{i1}...u_{in}A_i).$$

**Theorem 1.** ([18]) Let $X_1, ..., X_n, M$ be $n+1$ non-empty sets. Then $(EX_1...X_n M, \lor, \land)$ forms a completely distributive lattice under the binary compositions $\lor$ and $\land$ defined as follows. For any $\sum_{i \in I}(u_{i1}...u_{in}A_i), \sum_{j \in J}(v_{j1}...v_{jn}B_j) \in EX_1...X_n M$,

$$\sum_{i \in I}(u_{i1}...u_{in}A_i) \lor \sum_{j \in J}(v_{j1}...v_{jn}B_j) = \sum_{k \in I\cup J}(w_{i1}...w_{nk}C_k),$$

$$\sum_{i \in I}(u_{i1}...u_{in}A_i) \land \sum_{j \in J}(v_{j1}...v_{jn}B_j) = \sum_{i \in I, j \in J}[(u_{i1} \cap v_{j1} ... u_{in} \cap v_{jn})(A_i \cup B_j)],$$

where for any $k \in I \cup J$ (the disjoint union of $I$ and $J$), $C_k = A_k, w_{rk} = u_{rk}$ if $k \in I$, and $C_k = B_k, w_{rk} = v_{rk}$ if $k \in J, r = 1, 2, ..., n$.

$(EX_1...X_n M, \lor, \land)$ is called the $EI^{n+1}$ (expanding $n + 1$ sets $X_1, ..., X_n, M$) algebra over $X_1, ..., X_n$ and $M$. For $\alpha = \sum_{i \in I}(u_{i1}...u_{in}A_i), \beta = \sum_{j \in J}(v_{j1}...v_{jn}B_j) \in EX_1...X_n M$, $\alpha \leq \beta \iff \alpha \lor \beta = \beta \land (\forall u_{i1}...u_{in}A_i) (i \in I), (\forall v_{j1}...v_{jn}B_j) (h \in J)$ such that $A_i \supseteq B_h$, $u_{ri} \subseteq v_{rh}$, $1 \leq r \leq n$.

We first explain the $EI$ algebras (in the case $n = 0$ for the $EI^{n+1}$ algebras) using the semantic signification represented by the elements of $EM$. This enables us to “understand” the abstract AFS algebras intuitively, so our willingness to accept them is enhanced. Let $M$ be a set of attributes or concepts. The elements of $M$ are viewed as “elementary” (or “simple”) concepts. For $\sum_{i \in I} A_i \in EM$, the subset $A_i \subset C$ represents conjunction of the concepts in $A_i$, and $\sum_{i \in I} A_i$ is the disjunction of the conjunctions represented by $A_i$’s (i.e., every element of $EM$ corresponds to the disjunctive normal form of a formula representing a concept). For example, let $M = \{m_1, m_2, m_3\}$, where $m_1 = \text{Color red}, m_2 = \text{Color green}, m_3 = \text{Color blue}, m_4 = \text{Weight small}, m_5 = \text{Weight medium}, m_6 = \text{Weight large}, m_7 = \text{Textured yes} \text{ and } m_8 = \text{Textured no}$. Suppose the following fuzzy rules 1-4 describe a class of objects denoted by $C$:

**Rule 1:** If $x$ is $[\text{Color is blue}]$, $[\text{Textured is no}]$, then $x$ belongs to $C$. 

Rule 2: If $x$ is [Color is red], [Weight is small], then $x$ belongs to $C$.
Rule 3: If $x$ is [Color is red], [Weight is small], [Textured is no], then $x$ belongs to $C$.
Rule 4: If $x$ is [Color is red], [Weight is large], [Textured is yes], then $x$ belongs to $C$.

Rules 1-4 can be represented through a single rule by using the algebra operations $\lor$, $\land$ from $EM$. Define the symbol $\tau$ from $EM$ to the lattice representations of the membership degrees and fuzzy logic operations in lattice $H$.

Hence $\{m_3, m_8\} + \{m_1, m_4\} + \{m_1, m_6, m_7\}$ in lattice $EM$. This implies that the right side of the inequality is stricter than left side as the conditions of some rules. By (2) and (3), we observe that the operations $\lor$, $\land$ of the elements of $EM$ correspond to the “or”, “and” of the corresponding rules respectively.

In the following, we define an AFS structure, a triple $(M, \tau, X)$, which gives rise to the lattice representations of the membership degrees and fuzzy logic operations for each concept in $EM$. In general, where $X$ is universe of discourse and $M$ is a set of some simple concepts on $X$. The map $\tau : X \times X \rightarrow 2^M$ not only describes all binary relations corresponding to the (fuzzy or crisp) concepts in $M$, but also constructs a combinatorics object.

**Definition 5.** ([18,22]) Let $X, M$ be sets and $2^M$ be the power set of $M$, $\tau : X \times X \rightarrow 2^M$. $(M, \tau, X)$ is called an AFS structure if for any $x_1, x_2, x_3 \in X$, $\tau$ satisfies the following conditions: AX1: $\tau(x_1, x_2) \subseteq \tau(x_1, x_1)$, AX2: $\tau(x_1, x_2) \cap \tau(x_2, x_3) \subseteq \tau(x_1, x_3)$. $X$ is called universe of discourse, $M$ is called an attribute set and $\tau$ is called a structure.

In general, every concept in $M$ should be simple concept on $X$. Thus we can verify that $(M, \tau, X)$ is an AFS structure if $\tau$ is defined by

$$\tau(x_i, x_j) = \{m | m \in M, (x_i, x_j) \in R_m\}, x_i, x_j \in X.$$

**Theorem 2.** ([22]) Let $(M, \tau, X)$ be an AFS structure. For $x \in X$, $A \subseteq M$, we define the symbol

$$A(x) = \{y \mid y \in X, \tau(x, y) \supseteq A\}.$$

For any given $x \in X$, if we define $\phi_x : EM \rightarrow EXM$ as follows. For any $\sum_{i \in I} A_i \in EM, \phi_x(\sum_{i \in I} A_i) = \sum_{i \in I} A_i \cap \{x\})A_i \in EXM$, then $\phi_x$ is a lattice homomorphism from $(EM, \lor, \land)$ to $(EXM, \lor, \land)$. 

312 Y. REN, M. L. SONG, AND X. D. LIU
Theorem 2 implies that for any given concept \( \sum_{i \in I} A_i \in EM \), we get a map
\[
(\sum_{i \in I} A_i)(x) = \sum_{i \in I} A_i(\{x\})A_i \in EM.
\]
Since \((EM, \vee, \wedge)\) is a lattice, hence map \(\sum_{i \in I} A_i\) is a L-fuzzy set with membership degrees valued by \(EII\) algebra \(EM\). For \(\alpha, \beta \in EM\), \(\alpha \lor \beta\) and \(\alpha \land \beta\) are logic “or” and “and” of L-fuzzy sets \(\alpha\) and \(\beta\) respectively.

**Definition 6.** ([22]) Let \(X, M\) be sets, \((M, \tau, X)\) be an AFS structure and \((X, S, m)\) be a measurable space, where \(S\) is the \(\sigma\)-algebra on \(X\) and \(m\) be a finite and positive measure on \(S\) with \(0 < m(X) < \infty\). Then \((M, \tau, X, S, m)\) is called a semi-congnitive field. For fuzzy concept \(\gamma = \sum_{i \in I} A_i \in EM\), if \(\forall i \in I, \forall x \in X, A_i(\{x\}) \in S\) (refer to (5)), then fuzzy concept \(\sum_{i \in I} A_i\) is called measurable in the semi-congnitive field \((M, \tau, X, S, m)\) and its membership function is defined as follows:

\[
\mu_{\gamma}(x) = \sup_{1 \leq r \leq n}(m(A_i(\{X\}))/m(X)).
\]

For \(M\) a given set of simple concepts on universe of discourse \(X\), using the AFS structure \((M, \tau, X)\) and the AFS algebras, we get the membership functions and their fuzzy logic operations for the fuzzy concepts in \(EM\) by (6). The lattice value membership degrees of the fuzzy concepts and their logical operations are determined by the original data and facts, in stead of triangular norm and human intuition. This is totally different from the other fuzzy theories. We should notice that the membership functions of the fuzzy concepts in \(EM\) can be obtained as long as the binary relations of the simple concepts in \(M\) are given, and the real number descriptions of the objects on each attribute are not necessary. In [25], the authors defined the logic operator ‘ (negation) as follows: For fuzzy concept \(\sum_{i \in I} A_i \in EM\),
\[
(\sum_{i \in I} A_i)’ = \land_{i \in I}(\lor_{a \in A_i}\{a’\}),
\]
where \(a’\) is the negation of simple concept \(a \in M\) and the algorithm of obtaining \(a’\) can be found in [25]. The logic system \((EM, \lor, \land, ’)\) is called AFS fuzzy logic.

3. The fuzzy clustering analysis and the fuzzy clustering index

In this section, we study the fuzzy clustering algorithm and fuzzy clustering index proposed in [27]. Let \(X\) be the universe of discourse, \(M\) be a set of simple attributes on \(X\), \((M, \tau, X, S)\) be a semi-cognitive field established by the original data and facts. \(A \subseteq EM, A\) is a set of fuzzy sets which are selected to cluster the objects in \(X\). In [27], the authors gave the fuzzy clustering algorithm and fuzzy clustering index in the framework of AFS theory as followings:

**The algorithm of fuzzy clustering analysis based on AFS fuzzy logic**

**Step 1:**

Calculate fuzzy set \(\vartheta = \lor_{b \in A} b. x \in X\), the degrees of \(x\) belonging any cluster will be less than or equal to \(\mu_{\lor_{b \in A} b}(x)\). In general, each \(x\) should belong to \(\vartheta\) at a large degree. If for some \(x\), \(\mu_{\lor_{b \in A} b}(x)\) is too small, other fuzzy sets in \(EM\) should be selected and added into \(A\), in order to obtain a better clustering result.

**Step 2:**
For each \( x \in X \), find the fuzzy set \( \zeta_x \in (\Lambda)EI \), the fuzzy description of \( x \), where \((\Lambda)EI\) is the sub \( EI \) algebra generated by \( \Lambda \). \( \zeta_x \) should satisfy that \( \mu_{\zeta_x}(x) \) not only approaches to \( \mu_{\bigvee_{x \in X} b(x)} \), but also \( \mu_{\zeta_x}(y) \) is as small as possible for \( y \in X, y \neq x \). In other words, \( x \) can be distinguished by \( \zeta_x \) among other objects in \( X \) at maximum extent. For \( \varepsilon \geq 0 \) (in general \( \varepsilon \) is very small), define
\[
B^*_\varepsilon = \{ A_k \mid \mu_{A_k}(x) \geq \mu_{\bigvee_{x \in X} b(x)} - \varepsilon, k \in I, a = \sum_{i \in I} A_i \in \Lambda, a \text{ is irreducible} \},
\]
\[
\Lambda^\varepsilon = \{ \bigwedge_{\beta \in H} \beta \mid \mu_{\bigwedge_{\beta \in H} \beta}(x) \geq \mu_{\bigvee_{x \in X} b(x)} - \varepsilon \geq \mu_{\bigwedge_{\beta \in H} \beta}(x), \quad H \subseteq B^*_\varepsilon, \forall E \subseteq B^*_\varepsilon - H, E \neq \emptyset \},
\]
\( \zeta_x = \bigvee_{\alpha \in \Lambda^\varepsilon} \alpha \).

**Remark 1.** There are many alternative methods to find a fuzzy set \( \zeta_x \) in \((\Lambda)EI\) to describe each \( x \) \((x \in X)\) and control the rough extent of the fuzzy description. In order to distinguish \( x \) among other objects in \( X \) at maximum extent, \( \zeta_x \) represents the prototype of \( x \). Although it involves deep and difficult theory problems, human have applied it to deal with real world recognition problems. For example, human always give a description of the object using the chosen relative attributes and by the description of the object, other persons can find the similar objects in the set of objects and the similar objects are regarded as a cluster.

**Step 3:**

Apply \( \zeta_x \), the fuzzy description of each \( x \in X \) to establish the fuzzy relation matrix \( M = (m_{ij}) \) on \( X = \{ x_1, x_2, \ldots, x_n \} \), where
\[
(8) \quad m_{ij} = \min\{ \mu_{\zeta_{x_i} \wedge \zeta_{x_j}}(x_i), \mu_{\zeta_{x_i} \wedge \zeta_{x_j}}(x_j) \}.
\]

Theorem 3 in [27] ensures that there exists an integer \( r \) such that \((M^r)^2 = M^r\), i.e., fuzzy relation matrix \( Q = M^r \) can yield a partition tree with equivalence classes.

**Step 4:**

Let \( Q = M^r = (q_{ij}) \), the Boolean matrix \( Q_{\alpha} = (q_{ij}^\alpha) \), where \( q_{ij}^\alpha = 1 \Leftrightarrow q_{ij} \geq \alpha \), \( \alpha \in [0, 1] \). For \( \alpha \in [0, 1] \), \( x_i, x_j \in X, x_i, x_j \) are in the same cluster under threshold \( \alpha \) if and only if \( q_{ij}^\alpha = 1 \). For some \( x_i \in X \), if \( q_{ii}^\alpha = 0 \), then under fuzzy attribute set \( \Lambda \) for threshold \( \alpha \), which cluster \( x_i \) should belong to can not be determined.

**Step 5:**

Find the fuzzy descriptions \( \zeta_C \) for each cluster \( C \subseteq X \), where \( C \) is a cluster obtained in Step 4 under the threshold \( \alpha \), \( \zeta_C = \bigvee_{x \in C} \zeta_x \), the fuzzy description of class \( C \) whose membership degree \( \mu_{\zeta_C}(x) \) is not only the most approach to \( \mu_{\bigvee_{x \in X} b(x)} \), for each \( x \in C \), but also \( \mu_{\zeta_C}(y) \) is as small as possible for \( y \in X, y \notin C \). In other words, cluster \( C \) can be distinguished among other clusters in \( X \) at maximum extent. The fuzzy description of the boundary among the clusters \( C_1, C_2, \ldots, C_l \) is a fuzzy set \( \zeta_{bou} \in EM \),
\[
(9) \quad \zeta_{bou} = \bigvee_{1 \leq i, j \leq l, i \neq j} (\zeta_{C_i} \wedge \zeta_{C_j}),
\]
where \( \zeta_{C_i}, i = 1, 2, \ldots, l \), is the fuzzy description for the \( i \)th cluster.

**The fuzzy clustering index**

\( I_\alpha \), the evaluation of the fuzzy clustering results by the threshold \( \alpha \) is defined as follows:
\[
(10) \quad I_\alpha = \frac{\sum_{x \in \bigcup_{1 \leq i \leq l} C_i} \mu_{\zeta_{bou}}(x)}{\sum_{x \in \bigcup_{1 \leq i \leq l} C_i} \mu_{\zeta_{total}}(x)},
\]
where \( \zeta_{total} = \bigvee_{1 \leq i \leq l} \zeta_{C_i}, l \geq 2 \). The \( i \)th cluster is described by a fuzzy set \( \zeta_{C_i} \in EM \) which determines the degree of each object belonging to the \( i \)th cluster.
Fuzzy set \( \zeta_{bou} \) which describes the boundary among the classes shows the maximum degree of each object belonging to two different classes.

**Discussion:**

In the above clustering algorithm, \( \bigcup_{1 \leq i \leq l} C_i \) is the union of the clusters \( C_1, C_2, \ldots, C_l \) under threshold \( \alpha \in [0,1] \). We should notice that for \( x \in X \), if \( \mu_{\bigvee_{b \in \Lambda} b}(x) < \alpha \), then \( x \notin \bigcup_{1 \leq i \leq l} C_i \). But \( x \) also belongs to each fuzzy set \( \zeta_{C_i} \), \( i = 1, 2, \ldots, l \), at some degree. If \( \mu_{\zeta_{C_k}}(x) = \max_{1 \leq i \leq l}(\mu_{\zeta_{C_i}}(x)) \), then let \( x \) belong to cluster \( C_k \), i.e., \( x \in C_k \). This implies that although \( \mu_{\bigvee_{b \in \Lambda} b}(x) < \alpha \), we can determine which cluster \( x \) belongs by the fuzzy sets \( \zeta_{C_i} \), \( i = 1, 2, \ldots, l \). If for any \( x \in X - \bigcup_{1 \leq i \leq l} C_i \), \( \mu_{\bigvee_{b \in \Lambda} b}(x) > \alpha \), then let \( C_{l+1} = X - \bigcup_{1 \leq i \leq l} C_i \) be another cluster and \( \zeta_{C_{l+1}} = (\bigvee_{b \in \Lambda} b)' \) be the fuzzy description of cluster \( C_{l+1} \). Therefore, in any case, we can get a partition of \( X \). Let

\[
\zeta_{bou} = \bigvee_{1 \leq i,j \leq l, i \neq j}(\zeta_{C_i} \wedge \zeta_{C_j}), \zeta_{Total} = \bigvee_{1 \leq i \leq l} \zeta_{C_i},
\]

or

\[
\zeta_{bou} = \bigvee_{1 \leq i,j \leq l+1, i \neq j}(\zeta_{C_i} \wedge \zeta_{C_j}), \zeta_{Total} = \bigvee_{1 \leq i \leq l+1} \zeta_{C_i}.
\]

The fuzzy clustering index can be modified as follows:

\[
I_\alpha = \frac{\sum_{x \in X} \mu_{\zeta_{bou}}(x)}{\sum_{x \in X} \mu_{\zeta_{Total}}(x)}.
\]

The optimal threshold \( \alpha \) for the clustering analysis is \( \lambda = \min_{\alpha \in (0,1)} \{I_\alpha|Q_\alpha\} \) \( (Q_\alpha \) is defined in Step 4) and Boolean matrix \( Q_\lambda \) yields the best clustering for \( X \).

In step 2, the fuzzy description \( \zeta_x = \bigvee_{\alpha \in \Lambda_x} \alpha \), \( (\varepsilon > 0) \), which represents the main characteristic of \( x \) and determines the similarity relation between the objects in \( X \) and the clustering result. Parameter \( \varepsilon \) controls how many molecular elements are applied to describe the prototype of \( x \), i.e., \( \Lambda_x^\varepsilon \). Each element in \( \Lambda_x^\varepsilon \) represents one of the different characters for the prototype of \( x \). It is obvious that some characters may be more universal, e.g., \( \alpha \in \Lambda_x^\varepsilon \), more objects in \( X \) belong to \( \alpha \) at large degree, while some characters may be more individual, e.g., \( \alpha \in \Lambda_x^{\varepsilon} \), few objects in \( X \) belong to \( \alpha \) at large degree. By this observation, the appropriate characters, i.e., appropriate elements in \( \Lambda_x^{\varepsilon} \), can be selected to describe the prototype of \( x \). The universality and individuality of the fuzzy description of \( x \) can be controlled by the following algorithm:

\[
\Lambda_x^{\varepsilon,\delta} = \{ \alpha \in \Lambda_x^\varepsilon|c(\alpha) \leq \delta \}
\]

\[
\zeta_x^\delta = \bigvee_{\alpha \in \Lambda_x^{\varepsilon,\delta}} \alpha,
\]

Where \( \delta \in (0,1) \), for a fuzzy set \( \beta \in EM \), \( c(\beta) = \frac{\sum_{x \in X} \mu_{\beta}(x)}{|X|} \) is the cardinal ratio of \( \beta \) which measures the fuzzy cardinal number of fuzzy sets [1]. The elements in \( \Lambda_x^{\varepsilon,\delta} \) are the characters selected to describe the prototype of \( x \) whose universality and individuality are controlled by parameters \( \delta \). The larger \( \delta \), the more universal the fuzzy description of \( x \) is.
4. Experimental results

Example 1. We study the fuzzy clustering index defined in (10) and (13), and the universality and individuality of the fuzzy descriptions (14) by the Iris dataset. The well-known Iris dataset is provided by Fisher in 1936. The data have been obtained from the UCI ML repository [15]. The Iris data has 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 $X = \{x_1, x_2, ..., x_{150}\}$, where $x_i = (w_{i1}, w_{i2}, w_{i3}, w_{i4})$, $i = 1, 2, ..., 150$, be the set of the 150 samples and $M = \{m_1, m_2, ..., m_8\}$ be the set of concepts on the features petal length or width, where $m_1$ is the concept: the petal length is long, $m_2$ is the concept: the petal is not long, i.e., $m_2 = m_1'$, the negation of $m_1$, $m_3$ is the concept: the petal length is mid, i.e., close to $\sum_{i=1}^{150} w_{i3}/150 = 3.7587$, $m_4 = m_3'$, $m_5$ is the concept: the petal width is wide, $m_6 = m_4'$, $m_7$ is the concept: the petal width is mid, i.e. close to $\sum_{1\leq i\leq 150} w_{i4}/150 = 1.1987$, $m_8 = m_7'$. For each $m \in M$, according to the values of each pair $x_i$ and $x_j$ on the feature associating to concept $m$, the degrees of $x_i$ and $x_j$ belonging to concept $m$ can be compared, i.e., either $(x_i, x_j) \in R_m$ or $(x_j, x_i) \in R_m$ (refer to Definition 1). For example, if the petal length of $x_i$ is longer than or equal to that of $x_j$, then $(x_i, x_j) \in R_m$. Note that the petal lengths of $x_i$ and $x_j$ can be compared by human intuition, even if the exact petal lengths of $x_i$ and $x_j$ are not measured. In other words, to obtain the binary relation $R_m$ for a concept $m$, we only need the order relation description of the concept $m$. By Definition 3 in [27], one can verify that each $m \in M$ is a simple concept and for any $x, y \in X$, if define $\tau(x, y) = \{m|m \in M, (x, y) \in R_m\}$, then $(M, \tau, X)$ is an AFS structure. Let $S = 2^X$ be the $\sigma$-algebra on $X$ and $m$ be a measure on $S$ defined by: For any $A \subseteq X$, $m(A) = |A| (-A—$ is the number of elements of set $A)$. By Definition 6, one knows that $(M, \tau, X, S, m)$ is a semi-cognitive field and for any fuzzy concept $\gamma = \sum_{i \in J} A_i \in EM$, since $S = 2^X$, then $\gamma$ is measurable in the semi-cognitive field $(M, \tau, X, S, m)$ and its membership function is defined by (5).

Remark 2. $(M, \tau, X)$ is determined by the binary relations $R_m$, $m \in M$. In some situation, it is difficult or impossible to describe some features of objects using real numbers, considering the errors and noise with measures. For example, we do not describe hair white degree of a person by the number of white hair on his head. But the order relations can be easily and accurately established by the simple comparisons of each pair of objects. The order relations are enough to establish the AFS structure of a data system. The membership functions and their logic operations of the fuzzy concepts in $EM$ can be obtained by the AFS fuzzy logic system $(EM, \vee, \wedge, \prime)$ and the AFS structure $(M, \tau, X)$ (refer to (2)-(4) and (6),(7)).

The following Figures 1, 2 show the membership functions of fuzzy concepts $\{m_i\} \in EM$, $i = 1, 2, ..., 8$ defined by (6). $\{m_1\}$ to $\{m_4\}$ are the fuzzy concepts on feature petal length and $\{m_5\}$ to $\{m_8\}$ are on feature petal width. In figure 1, x-axe: the samples are numbered 1 to 150 by the petal length, i.e., the sample with longer petal is numbered larger number. In figure 2, x-axe: the samples are numbered 1 to 150 by the petal width, i.e., the sample with wider petal is numbered larger number. These membership functions are determined by the AFS structure $(M, \tau, X)$ which is dependent on the distribution of the original Iris data.

Let $\Lambda = \{\{m_1\}, \{m_2\}, ..., \{m_8\}\} \subseteq EM$, $\varepsilon > 0.1$. We study the clustering analysis by the fuzzy concepts in $\Lambda$. We set $\delta$ about 0.45 in (14). Thus there is
The samples are numbered 1 to 150 by the petal length membership degree. The membership functions of \( \{m_1\}, \{m_2\}, \{m_3\}, \{m_4\} \) on feature petal length.

Figure 1.

The membership functions of \( \{m_5\}, \{m_6\}, \{m_7\}, \{m_8\} \) on feature petal width.

Figure 2.

simply one molecular element in \( \Lambda^2_\varepsilon \), i.e. \( |\Lambda^2_\varepsilon| = 1 \), for each \( x \in X \). The fuzzy descriptions of some samples in \( X \) are listed as examples:

For \( x_1 \in X \), \( B^{0.1}_{x_1} = \{ \{m_2\}, \{m_4\}, \{m_6\} \} \), \( A^{0.1}_{x_1} = \{ \{m_2, m_4\}, \{m_2, m_6\} \} \), \( \zeta_{x_1} = \{m_2, m_4\} + \{m_2, m_6\} \), \( c(\{m_2, m_4\}) = 0.3002 \), \( c(\{m_2, m_6\}) = 0.4720 \), \( \zeta^{0.45}_{x_1} = \{m_2, m_4\} \).
For \( x_{13} \in X \), \( \Lambda_{x_{13}}^{0.1} = \{\{m_2, m_6\},\{m_6, m_8\}\} \), \( \zeta_{x_{13}} = \{m_2, m_6\} + \{m_6, m_8\} \), \( c(\{m_2, m_6\}) = 0.4720 \), \( c(\{m_6, m_8\}) = 0.3263 \), \( \zeta_{x_{13}}^{0.45} = \{m_6, m_8\} \).

For \( x_{53} \in X \), \( \Lambda_{x_{53}}^{0.1} = \{\{m_1\},\{m_3, m_7\}\} \), \( \zeta_{x_{53}} = \{m_1\} + \{m_3, m_7\} \), \( c(\{m_1\}) = 0.5194 \), \( c(\{m_3, m_7\}) = 0.4375 \), \( \zeta_{x_{53}}^{0.45} = \{m_3, m_7\} \).

For \( x_{57} \in X \), \( \Lambda_{x_{57}}^{0.1} = \{\{m_1\},\{m_5\},\{m_3, m_7\}\} \), \( \zeta_{x_{57}} = \{m_1\} + \{m_5\} + \{m_3, m_7\} \), \( c(\{m_1\}) = 0.5194 \), \( c(\{m_5\}) = 0.5385 \), \( c(\{m_3, m_7\}) = 0.4375 \), \( \zeta_{x_{57}}^{0.45} = \{m_3, m_7\} \).

For \( x_{124} \in X \), \( \Lambda_{x_{124}}^{0.1} = \{\{m_3\},\{m_1, m_5\}\} \), \( \zeta_{x_{124}} = \{m_3\} + \{m_1, m_5\} \), \( c(\{m_3\}) = 0.5195 \), \( c(\{m_1, m_5\}) = 0.4720 \), \( \zeta_{x_{124}}^{0.45} = \{m_1, m_5\} \).

For \( x_{133} \in X \), \( \Lambda_{x_{133}}^{0.1} = \{\{m_5, m_8\},\{m_1, m_5\}\} \), \( \zeta_{x_{133}} = \{m_5, m_8\} + \{m_1, m_5\} \), \( c(\{m_5, m_8\}) = 0.2900 \), \( c(\{m_1, m_5\}) = 0.4720 \), \( \zeta_{x_{133}}^{0.45} = \{m_5, m_8\} \).

The following Figures 3 to 8 show the membership functions of \( \zeta_x \) and \( \zeta_x^\delta \) for \( x_1, x_{53}, x_{124} \).

![Figure 3. Membership functions of fuzzy sets \( \zeta_{x_1} \) and \( \zeta_{x_1}^{0.45} \)](image)

By the figures, one can observe that for each \( x \in X \), the fuzzy description \( \zeta_x \) is rougher than the fuzzy description \( \zeta_x^\delta \) and the rough extent can be controlled by the parameter \( \delta \). For the Iris data, we find the clustering analysis results based on the fuzzy description \( \zeta_x \) \( x \in X \), are not so satisfactory, because the fuzzy description \( \zeta_x \) is too rough. In the following, the fuzzy description \( \zeta_x^\delta \), i.e., \( \delta = \min_{\beta \in \Lambda} \{c(\beta)\} \), is applied to describe \( x \in X \), and substituting \( \zeta_x \) in Step 3 (8) with \( \zeta_x^\delta \), we get the fuzzy equivalence relation matrix \( Q \), i.e., \( Q = M^r \), \( Q^2 = Q \) and some of its numbers show as follows:
Figure 4. Membership functions of fuzzy sets $\zeta_{x_{13}}$ and $\zeta_{x_{13}}^{0.45}$

Figure 5. Membership functions of fuzzy sets $\zeta_{x_{53}}$ and $\zeta_{x_{53}}^{0.45}$
In the following, we apply formula (11) to calculate fuzzy clustering index $I_\alpha$ in (13). Let $C_1, C_2, \ldots, C_l$ be the clusters by the fuzzy equivalence matrix $Q$ for the threshold $\alpha \in (0, 1)$ and $D_1, D_2, \ldots, D_q$ be the expected clusters. The accuracy rate is defined as
NEW APPROACHES TO THE FUZZY CLUSTERING VIA AFS THEORY 321

Figure 8. Membership functions of fuzzy sets $\zeta_{x_{133}}$ and $\zeta^{0.45}_{x_{133}}$

(15)

$$\frac{\sum_{1 \leq i \leq n} \max_{1 \leq k \leq q} |C_i \cap D_k|}{|X|}.$$  

Figure 9 shows the accuracy rate with the expected clusters: 1 to 50 iris-setosa, 51 to 100 iris-versicolor, 101 to 150 iris-virginica and the fuzzy clustering indexes $I_\alpha$ defined by (10) and (13).

By Figure 6, we know that $I_\alpha = 0.3750$ is the minimum value of $I_\alpha$, $\alpha \in (0.42, 0.6333)$ and the samples are clearly parted into two clusters: $C_1$: 1 to 50, $\zeta_{C_1} = \{m_6\} + \{m_2, m_4\}$, the fuzzy description of $C_1$ with the semantic interpretation: the iris whose petal width is not wide or whose petal length is not long and not mid. $C_2$: 51 to 150, $\zeta_{C_2} = \{m_1\} + \{m_3\} + \{m_5\} + \{m_7\}$, the fuzzy description of $C_2$ with the semantic interpretation: the iris whose petal length is long or mid, or whose petal width is wide or mid. Although the clusters by the fuzzy equivalence matrix $Q$ for the threshold $\alpha \in (0.42, 0.6333)$ are not the best accuracy rate refer to the expected clusters 1 to 50 iris-setosa, 51 to 100 iris-versicolor, 101 to 150 iris-virginica, one can observed by Figure 10 and Figure 11 that this is the clearest clustering by the petal length and width.

When $\alpha \in (0.6333, 0.65)$, $I_\alpha = 0.4797$, the samples are parted into three clusters $C_1, C_2, C_3$: $\zeta_{C_1} = \{m_6\} + \{m_2, m_4\}$, the fuzzy description of $C_1$ with the semantic interpretation: the iris whose petal width is not wide or whose petal length is not long and not mid. $\zeta_{C_2} = \{m_3\} + \{m_7\}$, the fuzzy description of $C_2$ with the semantic interpretation: the iris whose petal length is mid, or whose petal width mid, $\zeta_{C_3} = \{m_1\} + \{m_5\}$, the fuzzy description of $C_3$ with the semantic interpretation: the iris whose petal length is long, or whose petal width is wide. Figure 12 shows the membership functions of the fuzzy descriptions of the three clusters.

The samples are clustered by the fuzzy descriptions of the clusters $\zeta_{C_1}, \zeta_{C_2}, \zeta_{C_3}$, i.e., $x \in C_k$, if $\mu_{\zeta_{C_k}}(x) = \max_{1 \leq i \leq q} \{\mu_{\zeta_{C_i}}(x)\}$. 5 samples $x_{71}, x_{78}, x_{84}, x_{107},$...
Figure 9. The accuracy of the clustering under the threshold $\alpha$ and the old clustering index $I_\alpha$ defined by (10) and the new clustering index defined by (13).

Figure 10. The distribution of the samples for petal length and width.

$x_{120}$ are incorrectly clustered refer to the expected clusters 1 to 50 iris-setosa, 51 to 100 iris-versicolor, 101 to 150 iris-virginica. The accuracy rate is 96.67%. For the total four features of iris-data, the clustering accuracy rate is 89.33% by using the function kmeans in MATLAB toolbox, which is based on the well known k-mean clustering algorithm [3,4]. And for the the total four features of iris-data,
the clustering accuracy rate is also 89.33% by using the function fcm in MATLAB toolbox, which is based on the well known fuzzy c-mean clustering algorithm [10]. If the k-mean and fuzzy c-mean are applied to the last two features: petal length and width of iris-data, like our approaches in Example 1, then the clustering accuracy rate of k-mean is 96% and the clustering accuracy rate of fuzzy c-mean is 94.67%. We should notice that the k-mean and fuzzy c-mean algorithm can only be applied
to the datasets with numerical features and the number of clusters has to be given in advance. The examples further to show that the proposed fuzzy validation index is quite accurate to describe the clarity of the clustering results and it can be applied to determine the final clusters. Furthermore, the membership functions obtained by the fuzzy c-mean algorithm are just numerical functions which have not the semantic interpretations with the linguistic labels on the features.

Remark 3. The fuzzy clustering index $I_{\alpha}$ evaluates the clear or discriminate extent between the clusters obtained by the fuzzy clustering analysis with the threshold $\alpha \in (0, 1)$. The smaller $I_{\alpha}$, the clearer or more discriminate the clustering is.

5. Conclusion

In this paper, in the framework of AFS theory, we propose an algorithm to control the rough extent of fuzzy descriptions of objects for recognitions and the examples and the analysis show that the fuzzy clustering index $I_{\alpha}$ efficiently evaluates the clear or discriminate extent between the clusters obtained by the fuzzy clustering analysis with the threshold $\alpha \in (0, 1)$. The smaller $I_{\alpha}$, the clearer or more discriminate the clustering is. Furthermore, in the new fuzzy clustering analysis, the attributes of objects can be various data types or sub-preference relations (i.e., simple fuzzy concepts) while by other clustering algorithms, it is difficult or impossible to study clustering problems of these kinds. The design algorithm is comprehensible and similar to human recognition habit without model but the AFS fuzzy logic of the fuzzy sets in $EM$, instead of fuzzy clustering model, genetic algorithm and T-S fuzzy model. Because each fuzzy set $\zeta_{C_i} \in EM$ which represents a cluster has a definitive semantic interpretation in AFS fuzzy logic, the linguistic interpretation of the proposed fuzzy clustering algorithm is very comprehensible. Indeed, this approach also can be regard as the knowledge representation of the training data. This implies that the approaches in this paper are also a new data mining method. A very interested research topic is that find $\alpha \in (0, 1)$, $\varepsilon > 0$, $\delta \in (0, 1)$ (refer to (14)) to optimize the accurate rate of the clustering results we discussed in Section 3. We would like to share this new idea with more mathematicians, scientists and engineers.

References

NEW APPROACHES TO THE FUZZY CLUSTERING VIA AFS THEORY


Research Center of Information and Control, Dalian University of Technology, Dalian 116024, P. R. China

Department of Mathematics, Dalian Maritime University, Dalain 116026, P. R. China
E-mail: renyan1108@sohu.com, dxliuros@hotmail.com