The reduction and fusion of fuzzy covering systems based on the evidence theory

Tao Feng [a,b], Shao-Pu Zhang [c], Ju-Sheng Mi [a,*]

[a] College of Mathematics and Information Science, Hebei Normal University, Shijiazhuang 050016, China
[b] College of Science, Hebei University of Science and Technology, Shijiazhuang 050018, China
[c] Department of Mathematics and Physics, Shijiazhuang Tiedao University, Shijiazhuang 050043, China

ARTICLE INFO

ABSTRACT

This paper studies reduction of a fuzzy covering and fusion of multi-fuzzy covering systems based on the evidence theory and rough set theory. A novel pair of belief and plausibility functions is defined by employing a method of non-classical probability model and the approximation operators of a fuzzy covering. Then we study the reduction of a fuzzy covering based on the functions we presented. In the case of multiple information sources, we present a method of information fusion for multi-fuzzy covering systems, by which objects can be well classified in a fuzzy covering decision system. Finally, by using the method of maximum flow, we discuss under what conditions, fuzzy covering approximation operators can be induced by a fuzzy belief structure.

© 2011 Published by Elsevier Inc.

1. Introduction

The Dempster–Shafer theory of evidence is a method developed to model and manipulate uncertain, imprecise, incomplete, and even vague information. It was originated by Dempster's concept of lower and upper probabilities [6], and extended by Shafer [26] as a theory. The basic representational structure in this theory is a belief structure, which consists of a family of subsets, called focal elements, with associated individual positive weights summing to one. The fundamental numeric measures derived from the belief structure are a dual pair of belief and plausibility functions. Since its inception, evidential reasoning has emerged as a powerful methodology for pattern recognition, image analysis, diagnosis, knowledge discovery, information fusion, and decision making [37,41,48].

Another important method used to deal with uncertainty in information systems is the theory of rough sets [20,21]. As a mathematical method to deal with insufficient and incomplete data, it is a set-theory-based technique to handle data with granular structures by using two sets called the rough lower approximation and the rough upper approximation to approximate an object. By using this method, knowledge hidden in information systems may be revealed and expressed in the form of decision rules and its main idea is using the existing knowledge to approximate uncertain concepts and phenomena [21]. The theory of rough sets has obtained many achievements in both theoretical researches and application aspects. It provides some practical solutions for certain problems in information science, such as artificial intelligence, data mining, pattern recognition, knowledge discovery, knowledge representation and intelligent control. The classical definition of a *Pawlak* rough set is with reference to an equivalence relation. From both theoretical and practical viewpoints, the classical equivalence relation is a very stringent condition that may limit applications of rough sets. Various theories were therefore developed from an equivalence relation to more general mathematical concepts: algebraic

* Corresponding author.
  E-mail addresses: fengtao_new@163.com (T. Feng), shaopuzhang@hotmail.com (S.-P. Zhang), mijsh@263.net (J.-S. Mi).

methods of the theory of rough sets [2,16,46], a neighborhood system from topological space [27,28,45,54], a similarity relation and tolerance relation or arbitrary binary relation based on rough sets [30,33,53,56] and rough set theory has been successfully used for reducing redundant attributes, describing dependency among attributes, evaluating the significance of attributes, and dealing with inconsistent and incomplete data in knowledge and data analysis [12,19,31]. On the other hand, most of the knowledge in real life applications is fuzzy. Therefore, to promote *Pawlak* rough set model to fuzzy environment is a very natural problem. In fact, various fuzzy kinds of generalizations have been proposed in [5,8,9,18,22,35,47,50]. For example, Dubois and Prade firstly introduced the rough fuzzy set [8]. Alternatively, a fuzzy similarity relation can be used to replace an equivalence relation. They proposed the definition of fuzzy rough set in [9]. Meanwhile, the rough fuzzy set model may be used to deal with knowledge acquisition in information systems with fuzzy decisions [29]. Most types of the above-mentioned binary relation can be viewed as a covering or a fuzzy covering on the universe of discourse. So we pay more attention to the development of the covering and fuzzy covering based on rough set model. In [51], Zakowski gave covering-based rough set model. In [57], Zhu and Wang discussed the reduction for this model, and studied the axiomatic characterization of the lower approximation operator. Whereafter, several models of covering-based rough sets and comparison already appeared in literature [1,17,24,32–34,55,58]. Recently, Chen et al. [3] and Yang and Li [42] proposed a way to reduce the covering systems without decision attribute, which are databases characterized by coverings. Li and Yin [14] gave ways to knowledge reduction of covering decision systems based on information theory. Deng et al. [7], Feng et al. [10] and Li et al. [13] even established fuzzy rough set models based on coverings.

There are strong connections between rough set theory and Dempster–Shafer theory of evidence. It has been demonstrated that various belief structures are associated with various rough approximation spaces such that the different dual pairs of lower and upper approximation operators induced by the rough approximation spaces may be used to interpret the corresponding dual pairs of belief and plausibility functions induced by the belief structures [4,23,40,44,45]. The Dempster–Shafer theory of evidence may be used to analyze knowledge acquisition in information systems. It is well known that knowledge reduction is one of the hot research topics in rough set theory. Many authors studied attribute reduction based on the theory of evidence in various information systems, for example, in complete information systems [52], in incomplete information systems [15], in incomplete decision systems [36] and in ordered information systems [38] and if crisp set is replaced by fuzzy set, Yao et al. [43] discussed how to reduce the reflective fuzzy decision system by the belief function and the plausibility function. In most of the papers mentioned above, the probability assignment of various granules are the same or every element in the universe of discourse has the same probability. The reductions studied in these literatures are to maintain some approximation or probability estimate of decision classes unchanged. They did not consider the changes of the mass function and reduce the information system maintaining the mass function unchanged. If an information system is generated by a fuzzy covering, then the basic granules are also generated by the fuzzy covering and we think that probability assignment of a granule should be strongly related to the elements covered by the granule. It is clear that the probability space of the universe of discourse is not a classical probability model. So we attempt to propose a new mass function by employing the ratio of the elements covered by every granule in a fuzzy covering system. Then we use the corresponding belief function and the plausibility function to reduce the information system ensuring the probabilities of every element and mass function unchanged. Meanwhile, though there are a lot of papers studied the evidence theory combined with rough set theory, most of them concentrate on attribute reduction based on the evidence theory or generating belief and plausibility functions by lower and upper approximation operators. They did not discuss information fusion using the set of basic granules, or the set of all focal elements of the fusion mass function may not be the set of basic information granules. That is, the above information fusion is not based on rough set theory. Therefore, another motivation of this paper is how to fuse the multi-information systems based on rough set theory. If there are more than one fuzzy coverings, we should consider the multi-information fusion combining with some feature of multi-fuzzy covering systems to ensure the set of focal element being a normalized fuzzy covering of the universe of discourse. By the fused mass function, the new belief function and plausibility function can be obtained by the evidence theory and we investigate how to generate the fuzzy rough approximation operators by the new belief (plausibility) functions.

In this paper, we study fuzzy evidence theory based on fuzzy coverings. In Section 2, we review the fuzzy covering lower and upper approximations generated by a fuzzy covering in a finite universe of discourse. Section 3 gives a pair of belief and plausibility function and homologous mass function with respect to the lower and upper approximation operators based on a fuzzy covering. Then we discuss the reduction of a fuzzy covering by using the plausibility function in information systems and in decision tables respectively. In Section 4, we propose an information fusion method of multi-fuzzy coverings in fuzzy covering systems. We first define a new fusion mass function, and discuss the properties and applications of the corresponding belief and plausibility function. We then give lower and upper approximation operations generated by the belief function and plausibility function in two special cases: (1) the number of the focal elements in a fusion mass function is smaller than the number of the elements in U; (2) the number of the focal elements in a fusion mass function is bigger than the number of the elements in U, and the mass function value of the focal elements, assigned to the same object, in a fusion mass function are the same. Finally, in Section 5, we conclude the paper with a summary.

2. Basic concepts

2.1. Fuzzy covering approximation space

Let U be a finite and nonempty set called the universe of discourse. The class of all subsets (respectively, fuzzy subset) of U will be denoted by $\mathcal{P}(U)$ (respectively, by $\mathcal{F}(U)$). For any $A \in \mathcal{F}(U)$, the α-level and strong α-level of A will be denoted by $(A)_\alpha$ and $(A)_{\alpha+}$, respectively, that is, $(A)_\alpha = \{x \in U : A(x) \geq \alpha\}$ and $(A)_{\alpha+} = \{x \in U : A(x) > \alpha\}$, where $\alpha \in [0, 1]$, the unit interval, $(A)_0 = U$, and $(A)_{1+} = \emptyset$ and $(A \cup B)(x) = \max\{A(x), B(x)\}, (A \cap B)(x) = \min\{A(x), B(x)\}, \sim A(x) = 1 - A(x)$, $\forall A, B \in \mathcal{F}(U)$. A class of fuzzy sets $\mathcal{C} \subseteq \mathcal{F}(U)$ is called a fuzzy covering (introduced in [7]) of U, if $(1) \cup \{A | A \in \mathcal{C}\}(x) = 1$, $\forall x \in U$, $(2) \emptyset \notin \mathcal{C}$. If each fuzzy set A in fuzzy covering $\mathcal{C}$ is normalized, i.e. $A(x) = 1$ for at least one $x \in U$, then $\mathcal{C}$ is said to be normalized. For a fuzzy covering $\mathcal{C}, A \in \mathcal{C}$, if $A(x) = 1$, then x is covered by A.

Definition 2.1. Suppose U is a finite and nonempty universe of discourse, and $\mathcal{C} = \{C_1, C_2, \ldots, C_n\}$ is a fuzzy covering of U. For every $x \in U$, let $\mathcal{C}_x = \cap\{C_j : C_j \in \mathcal{C}, C_j(x) = 1\}$, then $Cov(C) = \{\mathcal{C}_x : x \in U\}$ is another fuzzy covering of U, which is called the fuzzy covering of U induced by $\mathcal{C}$.

For every $x \in U$, $\mathcal{C}_x$ is the minimal fuzzy set in $Cov(\mathcal{C})$ such that $\mathcal{C}_x(x) = 1$, that is, $\mathcal{C}_x$ is the most complete fuzzy description of x with respect to $\mathcal{C}$. For every $x, y \in U$, if $\mathcal{C}_x(y) = 1$ then $\mathcal{C}_x \supseteq \mathcal{C}_y$. Thus if $\mathcal{C}_x(y) = 1$ and $\mathcal{C}_y(x) = 1$, then $\mathcal{C}_x = \mathcal{C}_y$. Obviously, any element in $Cov(\mathcal{C})$ cannot be written as the union of other elements in $Cov(\mathcal{C})$. The core of a fuzzy set A is all the elements of U which covered by A: $(A)_1 = \{x \in U : A(x) = 1\}$. Obviously the cores of the fuzzy sets in a fuzzy covering form a covering. In the crisp case, $Cov(C)$ is a covering but may not be a partition.

Definition 2.2. $\forall \alpha \in [0, 1], (\mathcal{C}_x)_\alpha = \{y \in U : \mathcal{C}_x(y) \geq \alpha\}$, then $(\mathcal{C}_x)_\alpha$ is a crisp set in U. $\forall X \in \mathcal{P}(U)$, define:

$$\underline{\mathcal{C}}_\alpha(X) = \{x \in U : (\mathcal{C}_x)_\alpha \subseteq X\};$$

$$\overline{\mathcal{C}}_\alpha(X) = \{x \in U : (\mathcal{C}_x)_\alpha \cap X \neq \emptyset\}.$$

$\underline{\mathcal{C}}_\alpha(X)$ and $\overline{\mathcal{C}}_\alpha(X)$ are called α-level fuzzy covering lower and upper approximation of a crisp set X, respectively.

Definition 2.3. Suppose $\mathcal{C}$ is a fuzzy covering of U, $Cov(\mathcal{C})$ is the fuzzy covering induced by $\mathcal{C}$. $\forall A \in \mathcal{F}(U)$, define

$$\underline{\mathcal{C}}(A) = \bigvee_{\alpha \in [0,1]} (\hat{\alpha} \cap \underline{\mathcal{C}}_{1-\alpha}((A)_{\alpha+})); \quad \overline{\mathcal{C}}(A) = \bigvee_{\alpha \in [0,1]} (\hat{\alpha} \cap \overline{\mathcal{C}}_\alpha((A)_\alpha)),$$

where $\hat{\alpha}$ denotes the constant fuzzy set with its membership function $\hat{\alpha}(x) = \alpha$, $\forall x \in U$. Then $\underline{\mathcal{C}}(A)$ and $\overline{\mathcal{C}}(A)$ are called fuzzy covering lower and upper approximations of a fuzzy set A, respectively. $\underline{\mathcal{C}}$ and $\overline{\mathcal{C}}$ are referred to fuzzy covering lower and upper approximation operators, respectively. $(U, \mathcal{C})$ is a fuzzy covering approximation space.

It is easy to verify the following conclusions:

Theorem 2.1. *Suppose $\mathcal{C}$ is a fuzzy covering of U. The fuzzy covering upper and lower approximation operators satisfy the following properties: $\forall A, B \in \mathcal{F}(U), \alpha \in [0, 1]$,*

(1) $\overline{\mathcal{C}}(\emptyset) = \emptyset, \underline{\mathcal{C}}(U) = U$;

(2) $\underline{\mathcal{C}}(A) \subseteq A \subseteq \overline{\mathcal{C}}(A)$;

(3) $\underline{\mathcal{C}}(A) = \sim \overline{\mathcal{C}}(\sim A), \overline{\mathcal{C}}(A) = \sim \underline{\mathcal{C}}(\sim A)$;

(4) $\underline{\mathcal{C}}(A \cup \hat{\alpha}) = \underline{\mathcal{C}}(A) \vee \hat{\alpha}, \overline{\mathcal{C}}(A \cap \hat{\alpha}) = \overline{\mathcal{C}}(A) \wedge \hat{\alpha}$;

(5) $\underline{\mathcal{C}}(A \cap B) = \underline{\mathcal{C}}(A) \cap \underline{\mathcal{C}}(B), \overline{\mathcal{C}}(A \cup B) = \overline{\mathcal{C}}(A) \cup \overline{\mathcal{C}}(B)$;

(6) $A \subseteq B \Rightarrow \overline{\mathcal{C}}(A) \subseteq \overline{\mathcal{C}}(B), \underline{\mathcal{C}}(A) \subseteq \underline{\mathcal{C}}(B)$;

(7) $\underline{\mathcal{C}}(A \cup B) \supseteq \underline{\mathcal{C}}(A) \cup \underline{\mathcal{C}}(B), \overline{\mathcal{C}}(A \cap B) \subseteq \overline{\mathcal{C}}(A) \cap \overline{\mathcal{C}}(B)$.

Theorem 2.2. *Let U be a finite and nonempty universe, $\mathcal{C}$ be a fuzzy covering of U. The fuzzy covering upper and lower approximations satisfy the following equations: $\forall A, B \in \mathcal{F}(U), x \in U$,*

$$\underline{\mathcal{C}}(A)(x) = \bigwedge_{y \in U} \{[1 - \mathcal{C}_x(y)] \vee A(y)\}; \quad \overline{\mathcal{C}}(A)(x) = \bigvee_{y \in U} \{\mathcal{C}_x(y) \wedge A(y)\}.$$

Assume $1_{\{y\}}$ denotes the fuzzy singleton with value 1 at y and 0 elsewhere; 1_X denotes the characteristic function of X, $X \subseteq U$.

Proposition 2.3. *Suppose $\mathcal{C}$ is a fuzzy covering of U, then*

    (1) $\overline{\mathcal{C}}(1_y)(x) = C_x(y), \ \forall x, y \in U.$

    (2) $\underline{\mathcal{C}}(1_{U \setminus \{y\}})(x) = 1 - C_x(y), \ \forall x, y \in U.$

    (3) $\overline{\mathcal{C}}(1_X)(x) = \max\{C_x(y) \mid y \in X\}, \ \forall x \in U, X \in \mathcal{P}(U).$

    (4) $\underline{\mathcal{C}}(1_X)(x) = \min\{1 - C_x(y) \mid y \notin X\}, \ \forall x \in U, X \in \mathcal{P}(U).$

    (5) $\underline{\mathcal{C}}(\hat{\alpha}) = \overline{\mathcal{C}}(\hat{\alpha}) = \hat{\alpha}, \ \forall \alpha \in [0, 1].$

Proposition 2.4. *Suppose $\mathcal{C}$ is a fuzzy covering of U. $B \subseteq \mathcal{C}$ is also a fuzzy covering of U, then $\forall X \in \mathcal{F}(U), \forall x \in U$,*

    (1) $B_x \supseteq C_x$;

    (2) $\underline{B}(X)(x) \leq \underline{\mathcal{C}}(X)(x)$ and $\overline{B}(X)(x) \geq \overline{\mathcal{C}}(X)(x).$

2.2. Belief and plausibility functions

Basing on the concept of information granularity and the theory of possibility, Zadeh first generalized the Dempster–Shafer theory to fuzzy situation [49,50]. First of all, the belief structure should be generalized to fuzzy environment.

Definition 2.4 [37]**.** Let U be a nonempty finite set. A set function $m : \mathcal{F}(U) \rightarrow [0, 1]$ is referred to a basic probability assignment (also called mass function) if it satisfies axioms (M1) and (M2):

    (M1) $m(\emptyset) = 0;$ (M2) $\sum\limits_{A \in \mathcal{F}(U)} m(A) = 1.$

A fuzzy set $X \in \mathcal{F}(U)$ with $m(X) > 0$ is referred to a focal element of m. We denote by $\mathcal{M}$ the family of all focal elements of m. The pair $(\mathcal{M}, m)$ is called a fuzzy belief structure. In the following discussions, all the focal elements are supposed to be normalized, i.e., for any $A \in \mathcal{M}$, there exists an $x \in U$ such that $A(x) = 1$.

Based on a fuzzy belief structure $(\mathcal{M}, m)$ on a finite universe of discourse U, Zadeh [48] defined the expected certainty, denoted by $Bel(X)$, and the expected possibility, denoted by $Pl(X)$, as a dual pair of generalization of Dempster–Shafer belief and plausibility functions: for all $X \in F(U)$,

$$Bel(X) = \sum_{A \in \mathcal{M}} m(A) \inf(A \Rightarrow X); \qquad Pl(X) = \sum_{A \in \mathcal{M}} m(A) \sup(X \cap A),$$

where $\inf(A \Rightarrow X)$ measures the degree to which A is included in X and $\sup(X \cap A)$ measures the degree that X intersects with A. It is easy to verify that the expected certainty and the expected possibility degenerate into the crisp belief and plausibility functions when the belief structure $(\mathcal{M}, m)$ and X are crisp.

In what follows, let $(U, \mathcal{P}(U))$ be a measurable space, $P(A) = \sum\limits_{x \in U} A(x) P(x), \forall A \in \mathcal{F}(U)$, where $P(x) = P(\{x\})$ and $(U, \mathcal{P}(U), P)$ is a probability space. Wu et al. [37] studied the fuzzy belief and fuzzy plausibility functions in infinite universe of discourse.

Definition 2.5 [37]**.** Let U be a nonempty universe of discourse which may be infinite, and $\mathcal{I}$ an implicator on $[0, 1]$. For $A, B \in F(U)$, we define $I(A \subseteq B) = \wedge_{x \in U} \mathcal{I}(A(x), B(x)) = \wedge_{x \in U}(A \Rightarrow_{\mathcal{I}} B)(x)$, where $I(A \subseteq B)$ measures the degree to which A is included in X.

Definition 2.6 [37]**.** Let U be a nonempty universe of discourse which may be infinite, (M, m) a fuzzy belief structure on U, and $\mathcal{I}$ an implicator on $[0, 1]$. A fuzzy set function $Bel : \mathcal{F}(U) \rightarrow [0, 1]$ is referred to a generalized fuzzy belief function on U if for all $X \in \mathcal{F}(U)$,

$$Bel(X) = \sum_{A \in \mathcal{F}(U)} m(A) I(A \subseteq X) = \sum_{A \in \mathcal{F}(U)} m(A) \wedge_{x \in U} \mathcal{I}(A(x), X(x)).$$

The fuzzy set function $Pl : \mathcal{F}(U) \rightarrow [0, 1]$ is referred to a generalized fuzzy plausibility function on U:

$$Pl(X) = 1 - Bel(\sim X), \quad X \in \mathcal{F}(U).$$

Definition 2.7 [37]**.** If U is a countable set, P is a probability measure on U, W is a nonempty set which may be infinite, R is a serial fuzzy binary relation from U to W, then we call $((U, P), W, R)$ a fuzzy belief space.

Theorem 2.5 [37]. *Let $((U, P), W, R)$ be a fuzzy belief space, and $\mathcal{I}$ an left and right monotonic, semicontinuous, border implicator on [0,1]. $\underline{R_{\mathcal{I}}}$ and $\overline{R_{\mathcal{I}}}$ are the $\mathcal{I}$-fuzzy rough approximation operators defined as the following: $\forall X \in \mathcal{F}(W), \ x \in U$*

$$\underline{R_{\mathcal{I}}}(A)(x) = \wedge_{y \in U} \mathcal{I}(R(x, y), A(y)); \qquad \overline{R_{\mathcal{I}}}(A)(x) =\sim \underline{R_{\mathcal{I}}}(\sim A).$$

Denote $Bel(X) = P(\underline{R_{\mathcal{I}}}(X)), X \in F(W), Pl(X) = P(\overline{R_{\mathcal{I}}}(X)), X \in F(W)$. Then $Bel : F(W) \to [0, 1]$ and $Pl : F(W) \to [0, 1]$ are a dual pair of fuzzy belief and plausibility functions.

3. Reduction of a fuzzy covering based on the evidence theory

Reduction of an information system is an important issue in rough set theory. In this section, we discuss the reduction of a fuzzy covering based on the evidence theory.

3.1. Belief and plausibility functions induced by general fuzzy covering rough sets

Suppose $\mathcal{C}$ is a fuzzy covering on U. In the following, we assume $\cap \emptyset = U$, where $\emptyset$ denotes the set which does not contain any element of $\mathcal{F}(U)$. (Since $(\mathcal{F}(U), \subseteq)$ is a complete lattice with $\emptyset$ being the least element and U the maximum element). In most cases, the probability of various granules is not always the same. Especially in a fuzzy covering space $(U, \mathcal{C}), \{\mathcal{C}_x : x \in U\}$ is the basic granule. Every covering class can cover many elements in U and every element in U maybe covered by more than one covering classes, every covering class must be used many times for approximating an object. So we should use a method of non-classical probability model to define the probability of granules. It is well know that $\{\mathcal{B}_x : x \in U, \mathcal{B} \subseteq \mathcal{F}(U)\}$ is a fuzzy covering of U if and only if $\{(\mathcal{B}_x)_1 : x \in U\}$ is a covering of U. Since $\{(\mathcal{C}_x)_1 : x \in U\}$ is a covering of U, we can use $\{(\mathcal{C}_x)_1 : x \in U\}$ to define mass function.

Theorem 3.1. *Let $U = \{x_1, x_2, \ldots, x_n\}$ be a nonempty and finite universe of discourse, $\mathcal{C}$ a fuzzy covering of U. $\forall X \in \mathcal{F}(U)$, define*

$$m_{\mathcal{C}}(X) = \begin{cases} \frac{|(X)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|}, & X \in Cov(\mathcal{C}) = \{\mathcal{C}_{x_1}, \ldots, \mathcal{C}_{x_n}\}; \\ 0, & otherwise. \end{cases}$$

Then $m_{\mathcal{C}}$ is a mass function. Denote

$$Bel_{\mathcal{C}}(X) = \sum_{A \in Cov(\mathcal{C})} \frac{|(A)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|} \bigwedge_{y \in U} ((1 - A(y)) \vee X(y)), \qquad Pl_{\mathcal{C}}(X) = \sum_{A \in Cov(\mathcal{C})} \frac{|(A)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|} \bigvee_{y \in U} (A(y) \wedge X(y)).$$

$Bel_{\mathcal{C}}$ and $Pl_{\mathcal{C}}$ are belief and plausibility functions on U respectively.

Proof. Since $\sum_{X \in \mathcal{F}(U)} m_{\mathcal{C}}(X) = \sum_{X \in Cov(\mathcal{C})} \frac{|(X)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|} = 1$ and $m_{\mathcal{C}}(\emptyset) = 0$, then we know $m_{\mathcal{C}}$ is a mass function.

By Definition 2.6, if $\mathcal{I}$ be an S-implicator based on a t-conorm $\mathcal{S}$, $\mathcal{T}$ a t-norm dual to $\mathcal{S}$, and $\mathcal{I}(A(x), X(x)) = (1 - A(x)) \vee X(x)$, $\mathcal{T}(A(x), X(x)) = A(x) \wedge X(x), \forall A, X \in \mathcal{F}(U), \ x \in U$, then we have

$$\begin{aligned} Bel_{\mathcal{C}}(X) &= \sum_{A \in \mathcal{M}} m_{\mathcal{C}}(A) \wedge_{x \in U} \mathcal{I}(A(x), X(x)) \\ &= \sum_{A \in \mathcal{M}} m_{\mathcal{C}}(A) \bigwedge_{y \in U} ((1 - A(y)) \vee X(y)) \\ &= \sum_{A \in Cov(\mathcal{C})} m_{\mathcal{C}}(A) \bigwedge_{y \in U} ((1 - A(y)) \vee X(y)) \\ &= \sum_{A \in Cov(\mathcal{C})} \frac{|(A)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|} \bigwedge_{y \in U} ((1 - A(y)) \vee X(y)). \end{aligned}$$

Thus $Bel_{\mathcal{C}}$ is a belief function on U. $Pl_{\mathcal{C}}$ being a plausibility function can be proved similarly. $\square$

If $\mathcal{C}$ is a fuzzy covering of U, then $\{\mathcal{C}_x, x \in U\} = \mathcal{M}$.

Theorem 3.2. *Let U be a nonempty and finite universe of discourse, $\mathcal{C}$ a fuzzy covering of U. $Bel_{\mathcal{C}}$ and $Pl_{\mathcal{C}}$ satisfy the following statements: $\forall A \in \mathcal{F}(U)$,*

1. *$Bel_{\mathcal{C}}(\emptyset) = Pl_{\mathcal{C}}(\emptyset) = 0, Bel_{\mathcal{C}}(U) = Pl_{\mathcal{C}}(U) = 1$;*
2. *$Bel_{\mathcal{C}}(A) \leq Pl_{\mathcal{C}}(A)$;*
3. *$Bel_{\mathcal{C}}(A) + Bel_{\mathcal{C}}(\sim A) \leq 1$;*

4. *$Bel_\mathcal{C}$ and $Pl_\mathcal{C}$ are all monotone about A;*
5. *$Bel_\mathcal{C}(A) + Pl_\mathcal{C}(\sim A) = 1$;*
6. *$Bel_\mathcal{C}(A) = Pl_\mathcal{C}(A)$ if and only if $\underline{\mathcal{C}}(A) = A = \overline{\mathcal{C}}(A)$.*

Proof. It is trivial by Wu et al. [37]. □

Proposition 3.3. *Let $U = \{x_1, x_2, \ldots, x_n\}$ be a nonempty and finite universe of discourse, $\mathcal{C}$ a fuzzy covering of U. Then $\forall X \in \mathcal{F}(U)$,*

$$Bel_\mathcal{C}(X) = \sum_{A \in Cov(\mathcal{C})} \frac{|(A)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|} \bigwedge_{y \in U} ((1 - A(y)) \vee X(y)) = \sum_{x \in U} \underline{\mathcal{C}}(X)(x) P_\mathcal{C}(x),$$

$$Pl_\mathcal{C}(X) = \sum_{A \in Cov(\mathcal{C})} \frac{|(A)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|} \bigvee_{y \in U} (A(y) \wedge X(y)) = \sum_{x \in U} \overline{\mathcal{C}}(X)(x) P_\mathcal{C}(x),$$

where $P_\mathcal{C}(x) = \frac{|(\mathcal{C}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1| |\{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|}$.

Proof. $\forall X \in \mathcal{F}(U), x \in U$,

$$\sum_{x \in U} \underline{\mathcal{C}}(X)(x) P_\mathcal{C}(x) = \sum_{x \in U} \left(\bigwedge_{y \in U} \{[1 - \mathcal{C}_x(y)] \vee X(y)\} \right) \frac{|(\mathcal{C}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1| |\{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|}$$

$$= \sum_{A \in Cov(\mathcal{C})} |\{y \in U : \mathcal{C}_y = A\}| \left(\bigwedge_{y \in U} \{[1 - A(y)] \vee X(y)\} \right) \frac{|(A)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1| |\{y \in U : \mathcal{C}_y = A\}|}$$

$$= \sum_{A \in Cov(\mathcal{C})} \left(\bigwedge_{y \in U} \{[1 - A(y)] \vee X(y)\} \right) \frac{|(\mathcal{C}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|}$$

$$= Bel_\mathcal{C}(X).$$

Similarly, we can prove that $Pl_\mathcal{C}(X) = \sum_{x \in U} \overline{\mathcal{C}}(X)(x) P_\mathcal{C}(x)$.
Thus, we can find that the probability of every element of the domain is not necessarily the same. □

3.2. Reduction of a fuzzy covering

In this subsection, by using the plausibility function, we discuss the reduction of a fuzzy covering to maintain the mass function unchanged.

3.2.1. Reduction of a fuzzy covering without decision attribute
Firstly, we study the reduction of a fuzzy covering without decision attribute.

Definition 3.1. Suppose U is a finite and non-empty universe of discourse, $Cov(\mathcal{C})$ is an induced fuzzy covering of U by $\mathcal{C}$. $\mathcal{B} \subseteq \mathcal{C}$ is a subcovering of U, and $\mathcal{B}_x = \mathcal{C}_x, \forall x \in U$, then $\mathcal{B}$ is called a consistent set of $\mathcal{C}$. Furthermore, if $\forall \mathcal{B}' \subset \mathcal{B}$, $\mathcal{B}'$ is not a subcovering of U or $\exists x \in U$ such that $\mathcal{B}'_x \neq \mathcal{C}_x$, then $\mathcal{B}$ is a reduction of $\mathcal{C}$.

Theorem 3.4. *Let $(U, \mathcal{C})$ be a fuzzy covering information space and $\mathcal{B} \subseteq \mathcal{C}$ a subcovering. Then the following conditions are equivalent:*

(1) *$\mathcal{B}$ is a consistent set of $\mathcal{C}$.*

(2) *$\overline{\mathcal{B}}(1_x) = \overline{\mathcal{C}}(1_x), \forall x \in U$.*

(3) *$\underline{\mathcal{B}}(\sim 1_x) = \underline{\mathcal{C}}(\sim 1_x), \forall x \in U$.*

Proof. It is easy to prove by Definition 3.1. □

Proposition 3.5. *Let $(U, \mathcal{C})$ be a fuzzy covering information space and $\mathcal{B} \subseteq \mathcal{C}$. $\mathcal{B}$ is a consistent set of $\mathcal{C}$ then*

(1) $m_\mathcal{B}(X) = m_\mathcal{C}(X), \forall X \in \mathcal{F}(U)$.
(2) $P_\mathcal{B}(x) = P_\mathcal{C}(x), \forall x \in U$.
(3) $Pl_\mathcal{B}(X) = Pl_\mathcal{C}(X)$ and $Bel_\mathcal{B}(X) = Bel_\mathcal{C}(X), \forall X \in \mathcal{F}(U)$.

Proof. (1) Since $\mathcal{B}$ is a consistent set of $\mathcal{C}$, we know that $\mathcal{B}_x = \mathcal{C}_x$, $\forall x \in U$. Then $m_{\mathcal{C}}(X) = m_{\mathcal{B}}(X)$, $\forall X \in \mathcal{F}(U)$ (see Theorem 3.1).

(2) By $\mathcal{B}_x = \mathcal{C}_x$, $\forall x \in U$, we have $P_{\mathcal{C}}(x) = P_{\mathcal{B}}(x)$, $\forall x \in U$. So (2) holds.

(3) Using $\mathcal{B}_x = \mathcal{C}_x$, $\forall x \in U$ and the definition of fuzzy covering upper and lower approximation, we have $Pl_{\mathcal{B}}(X) = Pl_{\mathcal{C}}(X)$ and $Bel_{\mathcal{B}}(X) = Bel_{\mathcal{C}}(X)$, $\forall X \in \mathcal{F}(U)$.

Thus we know that a consistent set of $\mathcal{C}$ ensures not only the basic granules but also the probabilities of every element in U unchanged. □

Obviously, we can use the dependency degree to reduce the fuzzy covering. But the dependency degree does not reflect the probability distribution of the basic granules. So we define another measure using plausibility function to reduce the fuzzy covering. This measure should be closely related to basic granules and probability distribution of the universe of discourse.

Definition 3.2. Let $U = \{x_1, x_2, \ldots, x_n\}$ be a nonempty and finite universe of discourse, $\mathcal{C}$ a fuzzy covering of U. $\mathcal{B} \subseteq \mathcal{C}$, then define

$$\gamma(\mathcal{B}, \mathcal{C}) = \prod_{i=1}^{n} \frac{\sum_{x \in U} \left(\frac{|(\mathcal{B}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1||\{y \in U: \mathcal{C}_y = \mathcal{C}_x\}|} \right) \overline{\mathcal{B}}(1_{x_i})(x)}{Pl_{\mathcal{C}}(1_{x_i})}$$

as the closeness degree of $\mathcal{B}$ to $\mathcal{C}$.

Proposition 3.6. *Let $U = \{x_1, x_2, \ldots, x_n\}$ be a nonempty and finite universe of discourse, $\mathcal{C}$ a fuzzy covering of U. $\mathcal{B} \subseteq \mathcal{C}$ is a subcovering of U, then $\mathcal{B}$ is a reduction of the fuzzy covering $\mathcal{C}$ iff $\gamma(\mathcal{B}, \mathcal{C}) = 1$, and for any non-empty proper subset $\mathcal{B}' \subset \mathcal{B}$, $\gamma(\mathcal{B}', \mathcal{C}) > 1$ or $\mathcal{B}'$ is not a subcovering of U.*

Proof. If $\mathcal{B} \subseteq \mathcal{C}$ is a subcovering, then

$$\prod_{i=1}^{n} \frac{\sum_{x \in U} \frac{|(\mathcal{B}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1||\{y \in U: \mathcal{C}_y = \mathcal{C}_x\}|} \overline{\mathcal{B}}(1_{x_i})(x)}{Pl_{\mathcal{C}}(1_{x_i})} = 1 \quad \text{(by Proposition 3.3)}$$

$$\Leftrightarrow \prod_{i=1}^{n} \frac{\sum_{x \in U} \frac{|(\mathcal{B}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1||\{y \in U: \mathcal{C}_y = \mathcal{C}_x\}|} \overline{\mathcal{B}}(1_{x_i})(x)}{\sum_{x \in U} \frac{|(\mathcal{C}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1||\{y \in U: \mathcal{C}_y = \mathcal{C}_x\}|} \mathcal{C}_x(x_i)} = 1 \quad \text{(by Proposition 2.4)}$$

$$\Leftrightarrow \frac{\sum_{x \in U} \frac{|(\mathcal{B}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1||\{y \in U: \mathcal{C}_y = \mathcal{C}_x\}|} \overline{\mathcal{B}}(1_{x_i})(x)}{\sum_{x \in U} \frac{|(\mathcal{C}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1||\{y \in U: \mathcal{C}_y = \mathcal{C}_x\}|} \mathcal{C}_x(x_i)} = 1, \ \forall x_i \in U$$

$$\Leftrightarrow \sum_{x \in U} \frac{|(\mathcal{B}_x)_1| \mathcal{B}_x(x_i)}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1||\{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|} = \sum_{x \in U} \frac{|(\mathcal{C}_x)_1| \mathcal{C}_x(x_i)}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1||\{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|}, \ \forall x_i \in U$$

$$\Leftrightarrow |(\mathcal{B}_x)_1| = |(\mathcal{C}_x)_1| \text{ and } \mathcal{B}_x(x_i) = \mathcal{C}_x(x_i), \ \forall x, x_i \in U \quad \text{(by Proposition 2.4)}$$

$$\Leftrightarrow \mathcal{B}_x = \mathcal{C}_x, \ \forall x \in U.$$

Thus $\mathcal{B}$ is a fuzzy covering of U and a reduction of $\mathcal{C}$ iff $\gamma(\mathcal{B}, \mathcal{C}) = 1$ and for any $\mathcal{B}' \subset \mathcal{B}$, $\gamma(\mathcal{B}', \mathcal{C}) > 1$ or $\mathcal{B}'$ is not a subcovering of U. □

Example 3.1. Let $U = \{a, b, c\}$, $\mathcal{C} = \{C_1 = 1/a + 0/b + 0/c, \ C_2 = 0/a + 1/b + o.7/c, \ C_3 = 0/a + 0/b + 1/c, \ C_4 = 1/a + 0.3/b + 1/c\}$.

$$\mathcal{C}_a = 1/a + 0/b + 0/c, \quad \mathcal{C}_b = 0/a + 1/b + 0.7/c, \quad \mathcal{C}_c = 0/a + 0/b + 1/c.$$

$$Pl_{\mathcal{C}}(1_a) = \frac{1}{3}, \quad Pl_{\mathcal{C}}(1_b) = \frac{1}{3}, \quad Pl_{\mathcal{C}}(1_c) = \frac{1.7}{3}.$$

Let $\mathcal{B} = \mathcal{C} - \{C_4\}$, $\mathcal{B}_a = 1/a + 0/b + 0/c$, $\mathcal{B}_b = 0/a + 1/b + 0.7/c$, $\mathcal{B}_c = 0/a + 0/b + 1/c$. $\gamma(\mathcal{B}, \mathcal{C}) = 1$, and for every proper subset $\mathcal{B}'$ of $\mathcal{B}$, $\gamma(\mathcal{B}', \mathcal{C}) \neq 1$. Thus $\mathcal{B}$ is a reduction of $\mathcal{C}$.

Theorem 3.7. *Let* $(U, \mathcal{C})$ *be a fuzzy covering information system and* $\mathcal{B} \subseteq \mathcal{C}$ *a subcovering of U.*

(1) *If* $P_{\mathcal{C}}(x) \leq P_{\mathcal{B}}(x)$, $\forall x \in U$, *then* $\mathcal{B}$ *is a consistent set of* $\mathcal{C}$ *if and only if* $Pl_{\mathcal{B}}(X) = Pl_{\mathcal{C}}(X)$, $\forall X \in \mathcal{F}(U)$.

(2) *If* $\forall x, y \in U$, $\mathcal{C}_x = \mathcal{C}_y \Leftrightarrow \mathcal{B}_x = \mathcal{B}_y$ *and* $\forall x \in U$, $(\mathcal{B}_x)_1 = (\mathcal{C}_x)_1$, *then* $\mathcal{B}$ *is a consistent set of* $\mathcal{C}$ *if and only if*
$Pl_{\mathcal{B}}(X) = Pl_{\mathcal{C}}(X)$, $\forall X \in \mathcal{F}(U)$.

Proof. It is obvious that (1) hold, we only prove (2).
(2) Since $Pl_{\mathcal{B}}(X) = Pl_{\mathcal{C}}(X)$, $\forall X \in \mathcal{F}(U)$, we have $Pl_{\mathcal{B}}(1_x) = Pl_{\mathcal{C}}(1_x)$, $\forall x \in U$.
Thus, $\forall x \in U$,

$$\sum_{A \in Cov(\mathcal{C})} \frac{|(A)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|} \bigvee_{y \in U} (A(y) \wedge 1_x(y)) = \sum_{A \in Cov(\mathcal{B})} \frac{|(A)_1|}{\sum_{Y \in Cov(\mathcal{B})} |(Y)_1|} \bigvee_{y \in U} (A(y) \wedge 1_x(y)).$$

If $\forall x, y \in U$, $\mathcal{C}_x = \mathcal{C}_y \Leftrightarrow \mathcal{B}_x = \mathcal{B}_y$, and $\forall x \in U$, $(\mathcal{B}_x)_1 = (\mathcal{C}_x)_1$, then $\frac{|(\mathcal{C}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1|} = \frac{|(\mathcal{B}_x)_1|}{\sum_{Y \in Cov(\mathcal{B})} |(Y)_1|}$, $\forall x \in U$. Hence
$\overline{\mathcal{B}}(1_x) = \overline{\mathcal{C}}(1_x)$, that is, $\mathcal{B}$ is a consistent set of $\mathcal{C}$. The proof of the converse is obvious. $\square$

Example 3.2 (Following Example 3.1). $\mathcal{B}_a = 1/a + 0/b + 0/c$, $\mathcal{B}_b = 0/a + 1/b + 0.7/c$, $\mathcal{B}_c = 0/a + 0/b + 1/c$.
$Pl_{\mathcal{B}}(1_a) = \frac{1}{3} = Pl_{\mathcal{C}}(1_a)$, $Pl_{\mathcal{B}}(1_b) = \frac{1}{3} = Pl_{\mathcal{C}}(1_b)$, $Pl_{\mathcal{B}}(1_c) = \frac{1.7}{3} = Pl_{\mathcal{C}}(1_b)$, then $\mathcal{B}$ is a consistent set of $\mathcal{C}$.

3.2.2. Reduction of a fuzzy covering decision system

Let U be a non-empty and finite universe, $\mathcal{C}$ a fuzzy covering of U, and D a set of decision attributes, then $(U, \mathcal{C}, D)$ will be called a fuzzy covering decision system. It is discussed that the reduction of a reflexive fuzzy decision system only maintaining the upper approximation of every decision class unchanged in [43], and the probability assignment of every focal element is the same. In the following, we study attribute reduction of a fuzzy covering decision system in probability space keeping the upper approximation of every decision class and mass function of every basic information granule unchanged.

Definition 3.3. Let $(U, \mathcal{C}, D)$ be a fuzzy covering decision system, $U = \{x_1, \ldots, x_n\}$. $U/D = \{D_1, D_2, \ldots, D_r\}$ is a set of decision equivalence classes, $\mathcal{B} \subseteq \mathcal{C}$. If for every $D_i \in U/D$, $\overline{\mathcal{B}}(D_i) = \overline{\mathcal{C}}(D_i)$ and $|(\mathcal{B}_x)_1| = |(\mathcal{C}_x)_1|$, $\forall x \in U$, then $\mathcal{B}$ is called a relative consistent set of $\mathcal{C}$ about D. Furthermore, for any proper subset $\mathcal{B}' \subset \mathcal{B}$, $\exists D_i \in U/D$, $\overline{\mathcal{B}'}(D_i) \neq \overline{\mathcal{C}}(D_i)$ or $\exists x \in U$ such that $|(\mathcal{B}'_x)_1| \neq |(\mathcal{C}_x)_1|$, then we call $\mathcal{B}$ a relative reduction of $\mathcal{C}$.

We can obtain directly the following property.

Proposition 3.8. *Let* $(U, \mathcal{C}, D)$ *be a fuzzy covering decision system,* $U = \{x_1, \ldots, x_n\}$, $U/D = \{D_1, D_2, \ldots, D_r\}$, $\mathcal{B} \subseteq \mathcal{C}$. *If* $\mathcal{B}$ *is a relative consistent set of* $\mathcal{C}$, *then*

(1) $m_{\mathcal{C}}(X) = m_{\mathcal{B}}(X)$, $\forall X \in \mathcal{F}(U)$.
(2) $P_{\mathcal{C}}(x) = P_{\mathcal{B}}(x)$, $\forall x \in U$, *when* $\mathcal{C}_x = \mathcal{C}_y \Leftrightarrow \mathcal{B}_x = \mathcal{B}_y$, $\forall x, y \in U$.

Definition 3.4. Let $(U, \mathcal{C}, D)$ be a fuzzy covering decision system, $U = \{x_1, \ldots, x_n\}$, $U/D = \{D_1, D_2, \ldots, D_r\}$, $\mathcal{B} \subseteq \mathcal{C}$, define

$$\gamma(\mathcal{B}, D) = \prod_{i=1}^{r} \frac{\left(\sum_{x \in U} \frac{|(\mathcal{B}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1| |\{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|} \overline{\mathcal{B}}(D_i) \right)(x)}{\sum_{i=1}^{r} Pl_{\mathcal{C}}(D_i)}.$$

Then $\gamma(\mathcal{B}, D)$ is called the closeness degree of $\mathcal{B}$ to D.

Theorem 3.9. *Let* $(U, \mathcal{C}, D)$ *be a fuzzy covering decision system,* $U = \{x_1, \ldots, x_n\}$, $U/D = \{D_1, D_2, \ldots, D_r\}$, *then* $\mathcal{B} \subseteq \mathcal{C}$ *is a relative reduction of* $\mathcal{C}$ *iff* $\gamma(\mathcal{B}, D) = 1$, *and for any nonempty proper subset* $\mathcal{B}' \subset \mathcal{B}$, *we have* $\gamma(\mathcal{B}', D) > 1$.

Proof

$$\prod_{i=1}^{r} \frac{\left(\sum_{x \in U} \frac{|(\mathcal{B}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1| |\{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|} \overline{\mathcal{B}}(D_i) \right)(x)}{\sum_{i=1}^{r} Pl_{\mathcal{C}}(D_i)} = 1$$

$$\Leftrightarrow \prod_{i=1}^{r} \frac{\left(\sum_{x \in U} \frac{|(\mathcal{B}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1| |\{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|} \overline{\mathcal{B}}(D_i) \right)(x)}{\sum_{x \in U} \frac{|(\mathcal{C}_x)_1|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1| |\{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|} \overline{\mathcal{C}}(D_i)(x)} = 1 \quad \text{(by Proposition 2.4)}$$

$$\Leftrightarrow \sum_{x \in U} \frac{|(\mathcal{B}_x)_1| \overline{\mathcal{B}}(D_i)(x)|}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1| \{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|} = \sum_{x \in U} \frac{|(\mathcal{C}_x)_1| \overline{\mathcal{C}}(D_i)(x)}{\sum_{Y \in Cov(\mathcal{C})} |(Y)_1| |\{y \in U : \mathcal{C}_y = \mathcal{C}_x\}|}, \quad i = 1, 2, \ldots, r$$

$$\Leftrightarrow |(\mathcal{B}_x)_1| \overline{\mathcal{B}}(D_i)(x) = |(\mathcal{C}_x)_1| \overline{\mathcal{C}}(D_i)(x), \; \forall x \in U, i = 1, \ldots, r \; \text{(by Proposition 2.4)}$$

$$\Leftrightarrow \overline{\mathcal{B}}(D_i)(x) = \overline{\mathcal{C}}(D_i)(x) \text{ and } |(\mathcal{B}_x)_1| = |(\mathcal{C}_x)_1|, \; \forall x \in U.$$

So $\mathcal{B}$ is a reduction of $\mathcal{C}$ iff $\gamma(\mathcal{B}, D) = 1$ and for any $\mathcal{B}' \subset \mathcal{B}, \gamma(\mathcal{B}', D) > 1$. $\square$

The purpose of relative reduction of a fuzzy covering decision system is to find a minimal subset of a fuzzy covering to preserve the upper approximations of decision classes and the mass function unchanged.

Example 3.3. Let $U = \{a, b, c\}, \mathcal{C} = \{C_1 = 1/a + 0/b + 0/c, \; C_2 = 0/a + 1/b + o.7/c, \; C_3 = 0/a + 0/b + 1/c, \; C_4 = 1/a + 0.3/b + 1/c\}, U/R_D = \{\{a\}, \{b, c\}\}$.

$$\mathcal{C}_a = 1/a + 0/b + 0/c, \quad \mathcal{C}_b = 0/a + 1/b + 0.7/c, \mathcal{C}_c = 0/a + 0/b + 1/c.$$

$$Pl_{\mathcal{C}}(\{a\}) = \frac{1}{3}, \quad Pl_{\mathcal{C}}(\{b, c\}) = \frac{2}{3},$$

Let $\mathcal{B} = \mathcal{C} - \{C_4\}, \gamma(\mathcal{B}, D) = 1$, and for every subset $\mathcal{B}'$ of $\mathcal{B}, \gamma(\mathcal{B}', D) \neq 1$. Then $\mathcal{B}$ is a relative reduction of $\mathcal{C}$.

We consider the relation of the belief function (plausibility function) with respect to a fuzzy covering and a relative consistent set of the fuzzy covering. Then we can easily obtain the following conclusion.

Theorem 3.10. *Let $I = (U, \mathcal{C}, D)$ be a fuzzy covering decision system and $\mathcal{B} \subseteq \mathcal{C}$. Then we have*

(1) If $\mathcal{B}$ is a relative consistent set of $\mathcal{C}$ about D, and $\forall x, y \in U, \mathcal{C}_x = \mathcal{C}_y \Leftrightarrow \mathcal{B}_x = \mathcal{B}_y$, then $P_{\mathcal{B}}(x) = P_{\mathcal{C}}(x), \; \forall x \in U$.
(2) If $\forall x, y \in U, \mathcal{C}_x = \mathcal{C}_y \Leftrightarrow \mathcal{B}_x = \mathcal{B}_y$, and $\mathcal{B}$ is a relative consistent set of $\mathcal{C}$ about D, then $\underline{\mathcal{B}}(D_i) = \underline{\mathcal{C}}(D_i) \Leftrightarrow Bel_{\mathcal{B}}(D_i) = Bel_{\mathcal{C}}(D_i), \; \forall D_i \in U/D$.
(3) If $\forall x, y \in U, \mathcal{C}_x = \mathcal{C}_y \Leftrightarrow \mathcal{B}_x = \mathcal{B}_y$, and $\mathcal{B}$ is a relative consistent set of $\mathcal{C}$ about D, then $Pl_{\mathcal{B}}(D_i) = Pl_{\mathcal{C}}(D_i), \; \forall D_i \in U/D$.
(4) If $\forall x \in U, P_{\mathcal{B}}(x) = P_{\mathcal{C}}(x)$, then B is a relative consistent set of $\mathcal{C}$ about D if and only if $Pl_B(D_i) = Pl_{\mathcal{C}}(D_i), \; \forall D_i \in U/D$.

4. Information fusion and approximation in multi-fuzzy covering systems

In practical problems, we need to consider the information coming from multiple sources, thus we should study how to fuse the information systems. In this section, we extend Dempster evidence fusion function to multi-fuzzy covering systems.

4.1. Information fusion in multi-fuzzy covering systems

Definition 4.1. Suppose U is a nonempty and finite universe of discourse, $\triangle = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_n\}$ is a set of fuzzy coverings. Then we call $(U, \triangle)$ a multi-fuzzy covering system.

By Section 3, we can obtain a mass function m_i for each fuzzy covering $\mathcal{C}_i$. Thus we must consider how to get a new fusion mass function using $\{m_1, \ldots, m_n\}$ in order to get that the set of all focal element in fusion information is the set of all new granules generating a new approximation space.

Let $(U, \triangle)$ be a multi-fuzzy covering system, F_i the set of all focal elements of m_i with respect to $\mathcal{C}_i. B_i \in F_i$ denotes an arbitrary element B_i in F_i.

$$m_{\triangle}(A) = \begin{cases} \dfrac{\sum_{\cap_{i=1}^n \{B_i \in F_i\} = A} m_1(B_1) m_2(B_2) \ldots m_n(B_n)}{\sum_{(\cap_{i=1}^n \{B_i \in F_i\})_1 \neq \emptyset} m_1(B_1) m_2(B_2) \ldots m_n(B_n)}, & (A)_1 \neq \emptyset; \\ 0, & \text{otherwise.} \end{cases}$$

Then $m_{\triangle}$ is a mass function.

Thus the fuzzy information fusion rule is an improvement of Dempster rule, which is introduced in [25]. The influence of conflict evidence to every focal element is not considered, that is, the influence of $\sum m_{\triangle}(A), A \in \mathcal{F}(U)$, and $A \in \{\cap_{i=1}^n \{B_i \in F_i\} : (\cap_{i=1}^n \{B_i \in F_i\})_1 = \emptyset\}$ is not considered. But the conflict evidence has different relevance to every focal element of $m_{\triangle}$. So they can affect the fusion results. Thus, we want to integrate the influence into the fusion function. Now we propose a new method of evidence fusion which can be seen as improvements in conjunctive rules.

Definition 4.2. Suppose U is a finite and nonempty universe of discourse. $A, B \in \mathcal{F}(U)$, denote

$$
d(A, B) = \begin{cases} \frac{|(A)_1 \cap (B)_1|}{|(B)_1|}, & (B)_1 \neq \emptyset; \\ 1, & \text{otherwise.} \end{cases}
$$

Then $d(A, B)$ is called the inclusion degree on $\mathcal{F}(U)$.

The concept of inclusion degree is introduced in [39]. It is obvious that $d(A, B) \in [0, 1]$. Then we can use the inclusion degree to revise mass assignment, that is, use inclusion degree to distribute the conflict evidence.

Denote $T_\triangle = \{\bigcap_i \{B_i \in F_i\} : (\bigcap_i \{B_i \in F_i\})_1 \neq \emptyset, \ F_i \text{ is the set of focal elements of } m_i \text{ w.r.t. } \mathcal{C}_i \in \triangle\}$.

Example 4.1. Let $U = \{a, b, c\}$, $(\mathcal{C}_1)_a = \frac{1}{a} + \frac{1}{b} + \frac{0.5}{c}$, $(\mathcal{C}_1)_b = \frac{0.4}{a} + \frac{1}{b} + \frac{0.5}{c}$, $(\mathcal{C}_1)_c = \frac{0.4}{a} + \frac{1}{b} + \frac{1}{c}$. $(\mathcal{C}_2)_a = \frac{1}{a} + \frac{0.3}{b} + \frac{0.2}{c}$, $(\mathcal{C}_2)_b = \frac{0.5}{a} + \frac{1}{b} + \frac{1}{c}$, $(\mathcal{C}_2)_c = \frac{0.5}{a} + \frac{0.3}{b} + \frac{1}{c}$. $T_\triangle = \{\frac{1}{a} + \frac{0.3}{b} + \frac{0.2}{c}, \frac{0.5}{a} + \frac{1}{b} + \frac{0.5}{c}, \frac{0.4}{a} + \frac{1}{b} + \frac{0.5}{c}, \frac{0.4}{a} + \frac{1}{b} + \frac{1}{c}, \frac{0.4}{a} + \frac{0.3}{b} + \frac{1}{c}\} \supseteq \{\triangle_a = \frac{1}{a} + \frac{0.3}{b} + \frac{0.2}{c}, \triangle_b = \frac{0.4}{a} + \frac{1}{b} + \frac{0.5}{c}, \triangle_c = \frac{0.4}{a} + \frac{0.3}{b} + \frac{1}{c}\}$, where $\triangle_x = \bigcap\{(\mathcal{C}_i)_y : (\mathcal{C}_i)_y(x) = 1, y \in U, \mathcal{C}_i \in \triangle\}$.

By Example 4.1, we have $T_\triangle \supseteq \{\triangle_x : x \in U\}$ and $\forall x \in U, \exists A \in T_\triangle$ such that $A(x) = 1$. Thus $T_\triangle$ is a new normalized fuzzy covering of U.

Theorem 4.1. *Let U be a nonempty and finite universe, $(U, \triangle)$ a multi-fuzzy covering system. Define for each $A \in F(U)$,*

$$
m_\triangle^*(A) = \begin{cases} \sum_{\cap_{i=1}^n \{B_i \in F_i\} = A} \left(\prod_{i=1}^n m_i(B_i) \right) + \sum_{(\cap_{i=1}^n \{B_i \in F_i\})_1 = \emptyset} \left(\frac{\sum_{i=1}^n d(A, B_i)}{\sum_{A_j \in T_\triangle} \sum_{i=1}^n d(A_j, B_i)} \prod_{i=1}^n m_i(B_i) \right), & A \in T_\triangle; \\ 0, & \text{otherwise.} \end{cases}
$$

Then $m_\triangle^*$ is a mass function. We call it the fusion mass function.

Proof. (1) $m_\triangle^*(\emptyset) = 0$.

(2) $\forall A \in \mathcal{F}(U)$, if $A \in T_\triangle$, then $A \in \{\bigcap_i \{B_i \in F_i\} : (\bigcap_i \{B_i \in F_i\})_1 \neq \emptyset\}$. Conversely, if $X \in \{\bigcap_i \{B_i \in F_i\} : (\bigcap_i \{B_i \in F_i\})_1 \neq \emptyset\}$, then $\exists B_1 \in F_1, \ldots, B_n \in F_n, X = B_1 \cap B_2 \cap \ldots \cap B_n$ and $(X)_1 \neq \emptyset$. Thus

$$
\sum_{A \in \mathcal{F}(U)} m_\triangle^*(A) = \sum_{A_k \in T_\triangle} m_\triangle^*(A_k)
$$

$$
= \sum_{A_k \in T_\triangle} \left(\sum_{\underset{i=1}{\overset{n}{\cap}} \{B_i \in F_i\} = A_k} \left(\prod_{i=1}^n m_i(B_i) \right) + \sum_{(\cap_{i=1}^n \{B_i \in F_i\})_1 = \emptyset} \left(\frac{\sum_{i=1}^n d(A_k, B_i)}{\sum_{A_j \in T_\triangle} \sum_{i=1}^n d(A_j, B_i)} \prod_{i=1}^n m_i(B_i) \right) \right)
$$

$$
= \sum_{A_k \in T_\triangle} \left(\sum_{\underset{i=1}{\overset{n}{\cap}} \{B_i \in F_i\} = A_k} \left(\prod_{i=1}^n m_i(B_i) \right) \right) + \sum_{A_k \in T_\triangle} \left(\sum_{(\cap_{i=1}^n \{B_i \in F_i\})_1 = \emptyset} \left(\frac{\sum_{i=1}^n d(A_k, B_i)}{\sum_{A_j \in T_\triangle} \sum_{i=1}^n d(A_j, B_i)} \prod_{i=1}^n m_i(B_i) \right) \right)
$$

$$
= \sum_{A_k \in T_\triangle} \left(\sum_{\underset{i=1}{\overset{n}{\cap}} \{B_i \in F_i\} = A_k} \left(\prod_{i=1}^n m_i(B_i) \right) \right) + \sum_{(\cap_{i=1}^n \{B_i \in F_i\})_1 = \emptyset} \left(\prod_{i=1}^n m_i(B_i) \right)
$$

$$
= 1 - \left(\sum_{(\cap_{i=1}^n \{B_i \in F_i\})_1 = \emptyset} \left(\prod_{i=1}^n m_i(B_i) \right) - \sum_{(\cap_{i=1}^n \{B_i \in F_i\})_1 = \emptyset} \left(\prod_{i=1}^n m_i(B_i) \right) \right) = 1.
$$

Thus $m_\triangle^*$ is a mass function. $\square$

A fuzzy set $X \in \mathcal{F}(U)$ with $m_\triangle^*(X) > 0$ is referred to a focal element of $m_\triangle^*$. We denote by $\mathcal{M}_\triangle$ the family of all focal elements of $m_\triangle^*$. Then $\mathcal{M}_\triangle$ is a fuzzy covering of U and the elements in $\mathcal{M}_\triangle$ are the basic information granules.

Example 4.2. Let $U = \{a, b, c\}$, $(\mathcal{C}_1)_a = \frac{1}{a} + \frac{0.5}{b} + \frac{0}{c}$, $(\mathcal{C}_1)_b = \frac{0.1}{a} + \frac{1}{b} + \frac{1}{c}$, $(\mathcal{C}_1)_c = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}$. $(\mathcal{C}_2)_a = \frac{1}{a} + \frac{0.5}{b} + \frac{0}{c}$, $(\mathcal{C}_2)_b = \frac{0.1}{a} + \frac{1}{b} + \frac{0.7}{c}$, $(\mathcal{C}_2)_c = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}$.

$T_\triangle = \{(\mathcal{C}_1)_a,\ (\mathcal{C}_1)_b,\ (\mathcal{C}_2)_b,\ (\mathcal{C}_1)_c\}$ and $(\mathcal{C}_1)_a \cap (\mathcal{C}_2)_b = \emptyset$, $(\mathcal{C}_1)_b \cap (\mathcal{C}_2)_a = \emptyset$. Let $A = (\mathcal{C}_1)_a$, $B = (\mathcal{C}_1)_b$, $C = (\mathcal{C}_2)_b$, $D = (\mathcal{C}_1)_c$. Then

$$m_1(A) = \frac{1}{6}, \quad m_1(B) = \frac{1}{3}, \quad m_1(C) = 0, m_1(D) = \frac{1}{2}.$$

$$m_2(A) = \frac{1}{5}, \quad m_2(B) = 0, \quad m_2(C) = \frac{1}{5}, \quad m_1(D) = \frac{3}{5}.$$

Thus $m_\triangle(A) = \frac{7}{27}, m_\triangle(B) = \frac{6}{27}, m_\triangle(C) = \frac{5}{27}. m_\triangle(D) = \frac{9}{27}$.

However, $m_\triangle^*(A) = \frac{172}{675}, m_\triangle^*(B) = \frac{289}{1350}, m_\triangle^*(C) = \frac{127}{675}, m_\triangle^*(D) = \frac{463}{1360}$.

Then $m_\triangle(A) \geq m_\triangle^*(A), m_\triangle(B) \geq m_\triangle^*(B)$, but $m_\triangle(C) \leq m_\triangle^*(C), m_\triangle(U) \leq m_\triangle^*(U)$. $m_\triangle^*$ is a minor adjustment to $m_\triangle$.

By Example 4.2, we know that the probability assignment of C and U become greater, and the probability assignment of A and B become smaller, that is, the influence to C and U by $\{\bigcap_{i=1}^n \{B_i \in F_i\} : (\bigcap_{i=1}^n \{B_i \in F_i\})_1 = \emptyset\}$ is greater than A and B. From this aspect, $m_\triangle^*$ is more reasonable than $m_\triangle$. So we use $T_\triangle$ and $m_\triangle^*$ to define the belief function and plausibility function on U.

Definition 4.3. Let $U = \{x_1, x_2, \ldots, x_n\}$ be a nonempty and finite universe of discourse, $\triangle$ a multi-fuzzy covering system of U. $\forall X \in \mathcal{F}(U)$, denote

$$Bel_\triangle(X) = \sum_{A \in \mathcal{F}(U)} m_\triangle^*(A) \bigwedge_{y \in U} ((1 - A(y)) \vee X(y)), \qquad Pl_\triangle(X) = \sum_{A \in \mathcal{F}(U)} m_\triangle^*(A) \bigvee_{y \in U} (A(y) \wedge X(y)).$$

Then $Bel_\triangle$ and $Pl_\triangle$ are belief and plausibility functions on U, respectively.

Theorem 4.2. $Bel_\triangle$ and $Pl_\triangle$ satisfy the following statements: $\forall X \in \mathcal{F}(U)$,

1. $Bel_\triangle(\emptyset) = Pl_\triangle(\emptyset) = 0, Bel_\triangle(U) = Pl_\triangle(U) = 1$;
2. $Bel_\triangle(X) \leq Pl_\triangle(X)$;
3. $Bel_\triangle(X) + Bel_\triangle(\sim X) \leq 1$;
4. $Bel_\triangle$ and $Pl_\triangle$ are all monotone about X;
5. $Bel_\triangle(X) + Pl_\triangle(\sim X) = 1$.

Let U be a non-empty and finite universe, $\triangle$ be a multi-fuzzy covering system of U, and D be a set of decision attributes, then $(U, \triangle, D)$ is called a fuzzy covering decision system. Now we can compute the belief degree of every decision class by using operator $Bel_\triangle$. It will help us to classify the objects. So we give the decision rules in the following.

Suppose U is a nonempty and finite universe of discourse, D is a decision attribute set. $U/D = \{D_1, \ldots, D_r\}$ is a crisp partition of U. Define mathematical expectation $E(D) = \frac{\sum_{D_i \in U/D} Bel_\triangle(D_i)}{|U/D|}$, and standard deviation $\sigma(D) = \sqrt{\frac{\sum_{D_i \in U/D}(Bel_\triangle(D_i) - E(D))^2}{|U/D|}}$.

Assumption 4.1. Let U be a nonempty and finite universe of discourse, $\triangle$ a multi-fuzzy covering system of U. U/B and U/D are two decision partitions. Then we can use the following rules to decide which decision partition is better than the other.

(1) $\forall B_i \in U/B$, if $Bel(B_i) \geq Bel(D_j)$ for every $D_j \in U/D$ with $B_i \cap D_j \neq \emptyset$, then U/B is better than U/D.

(2) If (1) is not satisfied, then compute $E(D)$ and $E(B)$. If $E(D) > E(B)$, then the classification effects of U/D is better than U/B. If $E(B) > E(D)$, then the classification effects of U/B is better than U/D. If $E(B) = E(D)$, we consider $\sigma(D)$ and $\sigma(B)$, and we think that classification effects of the variance smaller one is better than the other. Otherwise, we believe the classification effects of the two decisions are both good.

Assumption 4.2. Let U be a nonempty and finite universe of discourse, $\triangle$ a multi-fuzzy covering system of U. D is a decision attribute and U/D is a partition of U. If the classification of the elements is known expect one element $x \in U$, we compute $Bel_\triangle(D_i \cup \{x\})$ for each $D_i \in U/D$. Select D^* satisfying $Bel_\triangle(D^* \cup \{x\}) = max\{Bel_\triangle(D_i \cup \{x\}) : D_i \in U/D\}$. Then we determine that x belongs to decision class D^*.

Example 4.3. Let $U = \{x_1, x_2, x_3, x_4\}$ be a set of four patients, $E = \{a_1, a_2, a_3, a_4\}$ be a set of four attributes. $a_1 =:$ Heat, $a_2 =:$ Cough, $a_3 =:$ headache, $a_4 =:$ URI. $D = \{1 =: commoncold, 2 =: influenza\}$. For attribute a_1 we have $\mathcal{C}_1$:
$\mathcal{C}_1 = \{1/x_1 + 0.7/x_2 + 1/x_3 + 0.4/x_4,\ 0.3/x_1 + 1/x_2 + 0.3/x_3 + 1/x_4\}$.
For attribute a_2 we have $\mathcal{C}_2$:

$C_2 = \{1/x_1 + 1/x_2 + 0.3/x_3 + 0.7/x_4, \ 0.8/x_1 + 0.7/x_2 + 1/x_3 + 1/x_4\}$.

For attribute a_3 we have C_3:

$C_3 = \{1/x_1 + 0.7/x_2 + 1/x_3 + 1/x_4, \ 0.4/x_1 + 1/x_2 + 1/x_3 + 0.7/x_4\}$.

For attribute a_4 we have C_4:

$C_4 = \{1/x_1 + 0.7/x_2 + 1/x_3 + 0.5/x_4, \ 0.4/x_1 + 1/x_2 + 1/x_3 + 1/x_4\}$.

Thus for C_1 we have:

$$C_{1x_1} = C_{1x_3} = 1/x_1 + 0.7/x_2 + 1/x_3 + 0.4/x_4,$$

$$C_{1x_2} = C_{1x_4} = 0.3/x_1 + 1/x_2 + 0.3/x_3 + 1/x_4,$$

$$m(C_{1x_1}) = m(C_{1x_3}) = \frac{1}{2}, \quad m(C_{1x_2}) = m(C_{1x_4}) = \frac{1}{2}.$$

For C_2 we have:

$$C_{2x_1} = C_{2x_2} = 1/x_1 + 1/x_2 + 0.3/x_3 + 0.7/x_4,$$

$$C_{2x_3} = C_{2x_4} = 0.8/x_1 + 0.7/x_2 + 1/x_3 + 1/x_4.$$

$$m(C_{2x_1}) = m(C_{2x_2}) = \frac{1}{2}, \quad m(C_{2x_2}) = m(C_{2x_4}) = \frac{1}{2}.$$

For C_3 we have:

$$C_{3x_1} = C_{3x_4} = 1/x_1 + 0.7/x_2 + 1/x_3 + 1/x_4,$$

$$C_{3x_2} = 0.4/x_1 + 1/x_2 + 1/x_3 + 0.7/x_4,$$

$$C_{3x_3} = 0.4/x_1 + 0.7/x_2 + 1/x_3 + 0.7/x_4.$$

$$m(C_{3x_1}) = m(C_{2x_4}) = \frac{1}{2}, \quad m(C_{2x_2}) = \frac{1}{6}, \quad m(C_{2x_3}) = \frac{1}{3}.$$

$T_\triangle = \{A_1 = 1/x_1 + 0.7/x_2 + 0.3/x_3 + 0.4/x_4, \ A_2 = 0.8/x_1 + 0.7/x_2 + 1/x_3 + 0.4/x_4, \ A_3 = 0.4/x_1 + 0.7/x_2 + 1/x_3 + 0.4/x_4, \ A_4 = 0.3/x_1 + 1/x_2 + 0.3/x_3 + 0.7/x_4, \ A_5 = 0.3/x_1 + 0.7/x_2 + 0.3/x_3 + 1/x_4\}$ and $m_\triangle(A_1) \approx 0.188$, $m_\triangle(A_2) \approx 0.246, m_\triangle(A_3) \approx 0.246, m_\triangle(A_4) \approx 0.130 \ m_\triangle(A_5) \approx 0.190$

(1) If there are two decisions $D = \{D_1 = \{x_1, x_2\}, D_2 = \{x_3, x_4\}\}, B = \{B_1 = \{x_1, x_3\}, B_2 = \{x_2, x_4\}\}$ given by two doctors, then we can compute $Bel_\triangle(D_1)$ and $Bel_\triangle(D_2)$ in two conditions. Moreover, we can determine the reasonable one of D_1 and D_2 by using the belief degree of D_1 and D_2. In the case of decision D,

$$Bel_\triangle(D_1) = 0.152, \quad Bel_\triangle(D_2) = 0.18.$$

The belief degree of D_1 is 0.152. The belief degree of D_2 is 0.18.

But if $B = \{B_1 = \{x_1, x_3\}, B_2 = \{x_2, x_4\}\}$, then $Bel_\triangle(B_1) = 0.204, Bel_\triangle(B_2) = 0.224$. Obviously, the belief degree of every set in B is bigger than D. Thus, we can believe that the classification effects of B is better than D.

(2) If we know $D_1 = \{x_1, x_2\}, D_2 = \{x_4\}$, but we do not know which decision class x_3 belongs to, we compute that $Bel_\triangle(\{x_1, x_2, x_3\}) = 0.4776$, and $Bel_\triangle(\{x_3, x_4\}\}) = 0.18$. So we think that x_3 belongs to D_1.

4.2. Induced fuzzy covering approximation operators

In Section 3, a fuzzy approximation space generates a serial lower (upper) approximations, we know that the lower (upper) approximation can induce the belief (plausibility) function. A nature problem is under what conditions, lower (upper) approximation operator can be generated from the $Bel(Pl)$ functions? And how to do it? In this subsection, we study these questions by using the theory of maximum flow [11]. The definitions about graph, network, maximum flow, etc, can be found in [11].

Let $|T_\triangle|$ denote the number of elements in $T_\triangle$, and $|T_\triangle| \neq \infty$. Let A_i in $T_\triangle$ denote a vertex a_i, $V_1 = \{a_i : A_i \in T_\triangle\}$, and x_j in U denote a vertex x_j, $V_2 = \{x_j : x_j \in U\}$. If $A_i(x_j) = 1$, then there is a directed edge from a_i to x_j. Denote $v_{ij} = (a_i, x_j)$.

Definition 4.4. A directed bipartite graph (or bigraph) is a directed graph whose vertices can be divided into two disjoint sets U and V such that every directed edge connects a vertex in U to one in V; that is, U and V are independent sets.

Proposition 4.3. *If $V = V_1 \cup V_2, E = \{e_{ij}\}$ denotes the set of directed edges, then $G(V, E)$ is a directed Bipartite graph.*

Since there must exist an element A_i in $T_\triangle$ such that $A_i(x_j) = 1, \forall x_j \in U$, we have the following conclusion.

Proposition 4.4. *There are at least one directed edge connected to $x_j, \forall x_j \in V_2$.*

In the following, we discuss the lower (upper) approximation operator generated by the *Bel(Pl)* functions in two cases.

Case 1. $|T_\triangle| \leq |U|$.

Add a new vertex s, and let s connect every vertex $a_i \in V_1$, we denote $e_i = (s, a_i)$ and add a new vertex t, let t connect every vertex $x_j \in V_2$, we denote $e_j = (x_j, t)$. Thus we can get a new graph G'. The capacity on each directed edge $e_i = (s, a_i)$ is $|U|$, the capacity on each directed edge $e_j = (x_j, t)$ is 1, and the capacity on each directed edge $e_{ij} = (a_i, x_j) \in E$ is 1. Then G' is a network. A flow on a graph is a nonnegative integer-valued function f from the set of directed edges to Z_+. The flow on directed edge e_{ij} is f_{ij}, the flow on directed edge e_i is f_{si}, the flow on directed edge e_j is f_{jt}. Now we can get a maximum flow on G'. Let F be a flow and $F(G)$ be the value of the flow F on G.

Proposition 4.5. *If G' is a network defined as above, $|T_\triangle| \leq n$, then the value of a maximum flow on G' is $F_{max}(G') = |U|$.*

By the method of maximum flow, the proof of Proposition 4.5 is obvious. By Proposition 4.5 we know that $f_{jt} = 1$, $\forall x_j \in V_2$. Thus for every $x_j \in V_2$ there exists only one vertex a_i such that $f_{ij} = 1$ and the following proposition holds.

Proposition 4.6. *Let G' be a network defined as above, $|T_\triangle| \leq n$. Then there exists a maximum flow F such that $F(G') = |U|$ and $f_{si} \geq 1, \forall a_i \in V_1$.*

Proof. Firstly, we give a feasible flow F' such that $f_{si} = 1$ for every $a_i \in V_1$ and $f_{jt} \leq 1$ for every $x_j \in U$. Thus there are at least $|U| - |V_1|$ edges $e_{ij} \in E$ such that $f_{ij} = 0$ but $c_{ij} = 1$. Secondly, we try to increase the flow f_{s1} for $a_1 \in V_1$ to get a new feasible flow. If f_{s1} can not be increased, then consider f_{s2}, and repeat the same action one by one. Finally, we always can get a maximum flow F such that $F(G') = |U|$ and $f_{si} \geq 1, \forall a_i \in V_1$. $\square$

Theorem 4.7. *Suppose F is a maximum flow of G' such that $F(G') = |U|$ and $f_{si} \geq 1, \forall a_i \in V_1$. If $f_{ij} = 1$ in F, then let $\triangle^{x_j} = \{A_i\}$ be a fuzzy neighborhood of x_j. Every $A_i \in T_\triangle$ is a fuzzy neighborhood of at least one elements in U and for every $x_j \in U$, there exists one neighborhood.*

Let $U = \{x_1, x_2, \ldots, x_n\}$ be a nonempty and finite universe of discourse, $\triangle$ a multi-fuzzy covering system of U. If $|T_\triangle| \leq n$, then $\forall X \in \mathcal{F}(U)$,

$$Bel_\triangle(X) = \sum_{A \in \mathcal{F}(U)} m^*_\triangle(A) \bigwedge_{y \in U}((1 - A(y)) \vee X(y)), \quad \text{(by Theorem 4.7)}$$

$$= \sum_{\triangle^x \in T_\triangle} |\{y \in U : \triangle^y = \triangle^x\}| \left(\bigwedge_{y \in U}((1 - \triangle^x(y)) \vee X(y)) \frac{m^*_\triangle(\triangle^x)}{|\{y \in U : \triangle^y = \triangle^x\}|} \right)$$

$$= \sum_{x \in U} \bigwedge_{y \in U}((1 - \triangle^x(y)) \vee X(y)) \frac{m^*_\triangle(\triangle^x)}{|\{y \in U : \triangle^y = \triangle^x\}|}.$$

Let $P(x) = \frac{m^*_\triangle(\triangle^x)}{|\{y \in U : \triangle^y = \triangle^x\}|}, \forall x \in U$, then the above equation is equivalent to

$$Bel_\triangle(X) = \sum_{x \in U} \bigwedge_{y \in U}((1 - \triangle^x(y)) \vee X(y))P(x).$$

Similarly, we can get $Pl_\triangle(X) = \sum_{x \in U} \bigvee_{y \in U}(\triangle^x(y) \wedge X(y))P(x)$.
Denote

$$\underline{Apr}_\triangle(X)(x) = \bigwedge_{y \in U}((1 - \triangle^x(y)) \vee X(y)),$$

$$\overline{Apr}_\triangle(X)(x) = \bigvee_{y \in U}(\triangle^x(y) \wedge X(y)).$$

Then $\underline{Apr}_\triangle$ and $\overline{Apr}_\triangle$ are called upper and lower approximation operators of $\triangle$ induced by the *Bel* function and the *Pl* function, respectively. Thus

$$Bel_\triangle(X) = \sum_{x \in U} \underline{Apr}_\triangle(X)(x)P(x),$$

$$Pl_\triangle(X) = \sum_{x \in U} \overline{Apr}_\triangle(X)(x)P(x).$$

The properties of lower and upper approximation operators induced by the belief function and the plausibility function are shown below.

Theorem 4.8. *Suppose* $\triangle$ *is a multi-fuzzy covering system of U. The lower and upper approximation operators induced by the belief function and the plausibility function satisfy the following properties:* $\forall A, B \in \mathcal{F}(U), \alpha \in [0, 1]$,

(1) $\overline{Apr}_{\triangle}(\emptyset) = \emptyset,\ \underline{Apr}_{\triangle}(U) = U$;

(2) $\underline{Apr}_{\triangle}(A) \subseteq A \subseteq \overline{Apr}_{\triangle}(A)$;

(3) $\underline{Apr}_{\triangle}(A) =\sim \overline{Apr}_{\triangle}(\sim A),\ \overline{Apr}_{\triangle}(A) =\sim \underline{Apr}_{\triangle}(\sim A)$;

(4) $\underline{Apr}_{\triangle}(A \cup \hat{\alpha}) = \underline{Apr}_{\triangle}(A) \vee \hat{\alpha},\ \overline{Apr}_{\triangle}(A \cap \hat{\alpha}) = \overline{Apr}_{\triangle}(A) \wedge \hat{\alpha}$;

(5) $\underline{Apr}_{\triangle}(A \cap B) = \underline{Apr}_{\triangle}(A) \cap \underline{Apr}_{\triangle}(B),\ \overline{Apr}_{\triangle}(A \cup B) = \overline{Apr}_{\triangle}(A) \cup \overline{Apr}_{\triangle}(B)$;

(6) $A \subseteq B \Rightarrow \overline{Apr}_{\triangle}(A) \subseteq \overline{Apr}_{\triangle}(B),\ \underline{Apr}_{\triangle}(A) \subseteq \underline{Apr}_{\triangle}(B)$;

(7) $\underline{Apr}_{\triangle}(A \cup B) \supseteq \underline{Apr}_{\triangle}(A) \cup \underline{Apr}_{\triangle}(B),\ \overline{Apr}_{\triangle}(A \cap B) \subseteq \overline{Apr}_{\triangle}(A) \cap \overline{Apr}_{\triangle}(B)$.

Proof. It is easy to prove by the above definition of $\underline{Apr}_{\triangle}$ and $\overline{Apr}_{\triangle}$. $\square$

Example 4.4. $U = \{a, b, c, d\}, T_{\triangle} = \{A_1 = 1/a + 0.5/b + 0.3/c + 1/d, A_2 = 0.1/a + 1/b + 1/c + 1/d, A_3 = 1/a + 0.5/b + 1/c + 0.2/d\}$. Thus we can obtain a network G' and a feasible flow on G' as follows:
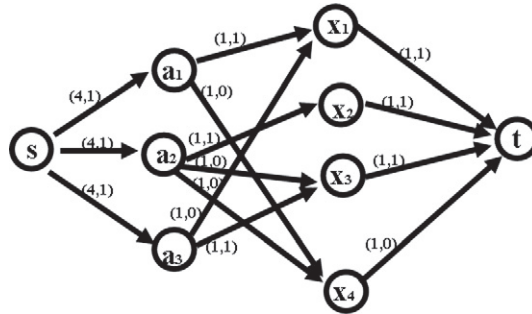
Now we can get a maximum flow
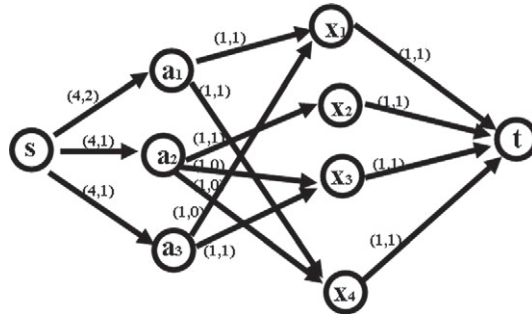
Then $\triangle^a = A_1, \triangle^b = A_2, \triangle^c = A_3, \triangle^d = A_1$. Thus we can compute the lower and upper approximations of every fuzzy set in U. If $X = 0.5/a + 0/b + 0.1/c + 1/d$, then $\underline{Apr}_{\triangle}(X) = 0.5/a + 0/b + 0.1/c + 0.5/d$ and $\overline{Apr}_{\triangle}(X) = 1/a + 1/b + 0.5/c + 1/d$.

Case 2. $|T_{\triangle}| > |U|$.

Similarly to Case 1, we assume that the flow capacity on each directed edge $e_i = (s, a_i)$ is 1, the capacity on each directed edge $e_j = (x_j, t)$ is $|T_{\triangle}|$, and the capacity on each directed edge $e_{ij} = (a_i, x_j) \in E$ is 1, which form G''. Hence we can get a maximum flow on G''.

Proposition 4.9. *If* G'' *is a network defined as above,* $|T_{\triangle}| > |U|$, *then the value of a maximum flow on* G'' *is* $F(G'') = |T_{\triangle}|$.

By the method of maximum flow, Proposition 4.9 is obvious. From Proposition 4.9, we know that $f_{jt} = 1, \forall a_i \in V_1$. Thus for every $a_i \in V_1$ there exists only one vertex x_j such that $f_{ij} = 1$. Then we obtain that if G'' is a network defined as above, $|T_{\triangle}| > |U|$, then there exists a maximum flow F such that $F(G'') = |T_{\triangle}|$ and $f_{jt} \geq 1, \forall x_j \in V_2$.

In the following we give the method of finding the maximum flow of G''.

Firstly, we give a feasible flow F' such that $f_{jt} = 1$ for every $x_j \in V_2$ and $f_{si} \leq 1$ for every $a_i \in V_1$. Thus there are at least $|T_\triangle| - |U|$ directed edges $e_{ij} \in E$ satisfying $f_{ij} = 0$, and then $f_{si} = 0$.

Secondly, we try to increase the flow f_{s1} for $a_1 \in V_1$ to get a new feasible flow such that $f_{jt} \geq 1, \forall x_j \in V_2$. If f_{s1} can not be increased, then consider f_{s2}, and repeat the same action one by one.

Finally, we can always get a maximum flow F such that $F(G'') = |T_\triangle|$ and $f_{jt} \geq 1, \forall x_j \in V_2$.

Theorem 4.10. *Suppose F is a maximum flow of G'' such that $F(G'') = |T_\triangle|$ and $f_{jt} \geq 1, \forall x_j \in V_2$. Let $\triangle^{x_j} = \{A_i : f_{ij} = 1\}$, $\forall x_j \in V_2$, then $\triangle^{x_j} \neq \emptyset, \forall x_j \in V_2$.*

Let $U = \{x_1, x_2, \ldots, x_n\}$ be a nonempty and finite universe of discourse, $\triangle$ a multi-fuzzy covering system of U. If $|T_\triangle| > n$ and for every $x \in U, m_\triangle^*(A_i) = m_\triangle^*(A_j), \forall A_i, A_j \in \triangle^x$, then $\forall X \in \mathcal{F}(U)$,

$$Bel_\triangle(X) = \sum_{A \in \mathcal{F}(U)} m_\triangle^*(A) \bigwedge_{y \in U} ((1 - A(y)) \vee X(y)) \quad (\text{by Theorem 4.10})$$

$$= \sum_{x \in U} \left(\sum_{A \in \triangle^x} \bigwedge_{y \in U} ((1 - A(y)) \vee X(y)) \right) m_\triangle^*(A)$$

$$= \sum_{x \in U} \frac{\sum_{A \in \triangle^x} \bigwedge_{y \in U} ((1 - A(y)) \vee X(y))}{|\triangle^x|} (m_\triangle^*(A)|\triangle^x|)$$

Let $P(x) = m_\triangle^*(A)|\triangle^x|$, then

$$Bel(X) = \sum_{x \in U} \frac{\sum_{A \in \triangle^x} \bigwedge_{y \in U} ((1 - A(y)) \vee X(y))}{|\triangle^x|} P(x)$$

Similarly, we get

$$Pl_\triangle(X) = \sum_{x \in U} \frac{\sum_{A \in \triangle^x} \bigvee_{y \in U} (A(y) \wedge X(y))}{|\triangle^x|} P(x).$$

Denote

$$\underline{Apr}'_\triangle(X)(x) = \frac{\sum_{A \in \triangle^x} \bigwedge_{y \in U} ((1 - A(y)) \vee X(y))}{|\triangle^x|}, \quad \overline{Apr}'_\triangle(X)(x) = \frac{\sum_{A \in \triangle^x} \bigvee_{y \in U} (A(y) \wedge X(y))}{|\triangle^x|}.$$

Then $\underline{Apr}'_\triangle$ and $\overline{Apr}'_\triangle$ are called upper and lower approximation operators of $\triangle$ induced by Bel and Pl, respectively. Thus $Bel_\triangle(X) = \sum_{x \in U} \underline{Apr}'_\triangle(X)(x)P(x), Pl_\triangle(X) = \sum_{x \in U} \overline{Apr}'_\triangle(X)(x)P(x)$.

The properties of lower and upper approximation operators induced by the belief function and the plausibility function are shown below.

Theorem 4.11. *Suppose $\triangle$ is a multi-fuzzy covering system of U. The upper and lower approximation operators induced by Bel and Pl satisfy the following properties: $\forall A, B \in \mathcal{F}(U)$,*

(1) $\overline{Apr}_\triangle(\emptyset) = \emptyset, \underline{Apr}_\triangle(U) = U$;

(2) $\underline{Apr}_\triangle(A) \subseteq A \subseteq \overline{Apr}_\triangle(A)$;

(3) $\underline{Apr}_\triangle(A) = \sim \overline{Apr}_\triangle(\sim A), \overline{Apr}_\triangle(A) = \sim \underline{Apr}_\triangle(\sim A)$;

(4) $A \subseteq B \Rightarrow \overline{Apr}_\triangle(A) \subseteq \overline{Apr}_\triangle(B), \underline{Apr}_\triangle(A) \subseteq \underline{Apr}_\triangle(B)$.

Example 4.5 (Following Example 4.3). By $|T_\triangle| = 5 > |U|$, we have $\triangle^{x_1} = \{A_1\}, \triangle^{x_2} = \{A_4\}, \triangle^{x_3} = \{A_2, A_3\}, \triangle^{x_4} = \{A_5\}$, and $|\triangle^3| = 2$. If $X = 1/x_1 + 1/x_2 + 0.3/x_3 + 0/x_4$, then $\underline{Apr}_\triangle(X) = 0.6/x_1 + 0.3/x_2 + 0.3/x_3 + 0/x_4, \overline{Apr}_\triangle(X) = 1/x_1 + 1/x_2 + 0.75/x_3 + 0.7/x_4$.

5. Conclusion

In this paper we first gave the definition and properties of a dual of fuzzy covering upper and lower approximation operators. Then we defined a new pair of belief function and plausibility function based on the fuzzy covering upper and lower approximation operators, and discussed their properties based on fuzzy coverings of a universe and then, we studied

the reduction of a fuzzy covering on the belief and plausibility functions and presented a method to compute a reduction by use of plausibility functions. Moreover, we proposed a fusion mass function over a multi-fuzzy covering system and discussed the application of the fusion mass function. Finally we discussed the question how to get lower and upper approximations in a evidence theory space of information fusion in two special cases. In the future, we will develop the proposed approaches to more generalized and more complex information systems such as fuzzy decision systems.

Acknowledgements

References

[1] D. Binaucci, G. Cattaneo, Information entropy and granular co-entropy of partitions and coverings: a summary, transactions on rough sets, Lecture Notes in Computer Science 5656 (2009) 15–66.
[2] Z. Bonikowksi, Algebraic structrues of rough sets, in: W.P. Ziarko (Ed.), Rough Sets, Fuzzy Sets and Knowledge Discovery, Springer-Verlag, Berlin, 1995, pp. 242–247.
[3] D.G. Chen, C.Z. Wang, Q.H. Hu, A new approach to attribute reduction of consistent and inconsistent covering decision systems with covering rough sets, Information Sciences 177 (2007) 3500–3518.
[4] D.G. Chen, W.X. Yang, F.C. Li, Measures of general fuzzy rough sets on a probabilistic space, Information Sciences 178 (2008) 3177–3187.
[5] M. De Cock, C. Cornelis, E.E. Kerre, Fuzzy rough sets: beyond the obvious, in: Proceedings of the 2004 IEEE International Conference on Fuzzy Systems, vol. 1, 2004, pp. 103–108.
[6] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, Annals of Mathematical Statistics 38 (1967) 325–339.
[7] T.Q. Deng, Y.M. Chen, W.L. Xu, Q.H. Dai, A novel approach to fuzzy rough sets based on a fuzzy covering, Information Sciences 177 (2007) 2308–2326.
[8] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems 17 (1990) 191–208.
[9] D. Dubois, H. Prade, Putting rough sets and fuzzy sets together, in: R. Slowinski (Ed.), Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, Dordrecht, 1992, pp. 203–232.
[10] T. Feng, J.S. Mi, W.Z. Wu, Covering-based generalized rough fuzzy sets, in: RSKT 2006, LNAI, vol. 4062, 2006, pp. 208–215.
[11] A.V. Goldberg, R.E. Tarjan, A new approach to the maximum-flow problem, Journal of the ACM 35 (4) (1988) 921–940., ACM Press.
[12] M. Kryszkiewicz, Rough set approach to incomplete information systems, Information Sciences 112 (1998) 39–49.
[13] T.J. Li, Y. Leung, W.X. Zhang, Generalized fuzzy rough approximation operators based on fuzzy coverings, International Journal of Approximate Reasoning 48 (2008) 836–856.
[14] F. Li, Y.Q. Yin, Approaches to knowledge reduction of covering decision systems based on information theory, Information Sciences 179 (2009) 1694–1704.
[15] P.J. Lingras, Y.Y. Yao, Data mining using extensions of the rough set model, Journal of the American Society for Information Science 49 (1998) 415–422.
[16] G.L. Liu, W. Zhu, The algebraic structures of generalized rough set theory, Information Sciences 178 (2008) 4105–4113.
[17] G.L. Liu, Y. Sai, A comparison of two types of rough se ts induced by coverings, International Journal of approximation Reasoning 50 (2009) 521–528.
[18] N.N. Moris, M.M. Yakout, Axiomatics for fuzzy rough sets, Fuzzy Sets and Systems 100 (1998) 327–342.
[19] E. Orlowska, Reasoning with incomplete information: rough set based information systems, Incompleteness and Uncertainty in Information Systems, Springer-Verlag, 1994, pp. 16–33.
[20] Z. Pawlak, Rough sets, International Journal of Computer Sciences 11 (1982) 341–356.
[21] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, Kluwer Academic Publishers, boston, 1991.
[22] Z. Pawlak, Rough sets and fuzzy sets, Fuzzy Sets and Systems 17 (1985) 99–102.
[23] C. Römer, A. Kandel, Applicability analysis of fuzzy inference by means of generalized Dempster-Shafer theory, IEEE Transactions on Fuzzy Systems 3 (1995) 448–453.
[24] P. Samanta, M.K. Chakraborty, Covering based approaches to rough sets and implication lattices, RSFDGrC 2009, 2009, pp. 127–134.
[25] K. Sentz, S. Ferson, , pp. 3–96. Combination of Evidence in Dempster–Shafer Theory, vol. 0835, Sandia National Laboratories, SAND, 2002..
[26] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.
[27] A. Skowron, J. Stepaniuk, Tolerance approximation spaces, Fundamenta Informaticae 27 (1996) 245–253.
[28] A. Skowron, J.F. Peters, Rough sets: trends and challenges, in: RSFDGrC 2003, LNAI, 2003, pp. 25–34.
[29] R. Slowinski, J. Stefanowski, Medical information systems-problems with analysis and way of solution, in: S.K. Pal, A. Skowron (Eds.), Rough Fuzzy Hybridiza-tion: A New Trend in Decision-making, Springer-Verlag, Singapore, 1999, pp. 301–315.
[30] R. Slowinski, D. Vanderpooten, A generalized definition of rough approximations based on similarity, IEEE Transaction on Knowledge and Data Engineering 12 (2000) 331–336.
[31] S. Trabelsi, S. Elouedi, Heuristic method for attribute selection from partially uncertain date using rough sets, International Journal of General System 39 (2010) 271–290.
[32] E. Tsang, D. Cheng, J. Lee, D. Yeung, On the upper approximations of covering generalized rough sets, in: Proceedings of the 3rd International Conference Machine Learning and Cybernetics, 2004, pp. 4200–4203.
[33] C.Z. Wang, C.X. Wu, D.G. Chen, A systematic study on attribute reduction with rough sets based on general binary relations, Information Sciences 178 (2008) 2237–2261.
[34] C.Z. Wang, D.G. Chen, C. Wu, Q.H. Hu, Data compression with homomorphism in covering information systems, International Journal of Approximation Reasoning 52 (2011) 519–525.
[35] W.Z. Wu, J.S. Mi, W.X. Zhang, Generalized fuzzy rough sets, Information Sciences 151 (2003) 263–282.
[36] W.Z. Wu, Attribute reduction based on evidence theory in incomplete decision systems, Information Sciences 178 (2008) 1355–1371.
[37] W.Z. Wu, Y. Leung, J.S. Mi, On generalized fuzzy belief functions in infinite spaces, IEEE Transactions on Fuzzy Systems 17 (2009) 385–397.
[38] W.H. Xu, X.Y. Zhang, J.M. Zhong, W.X. Zhang, Attribute reduction in ordered information systems based on evidence theory, Knowledge Information System 25 (2010) 169–184.
[39] Z.B. Xu, J.Y. Liang, C.Y. Dang, K.S. Chin, Inclusion degree: a perspective on measures for rough set data analysis, Information Sciences 141 (2002) 227–236.
[40] R.R. Yager, On the normalization of fuzzy belief structure, International Journal of Approximate Reasoning 14 (1996) 127–153.
[41] S.L. Yang, H. Luo, X.J. Hu, A combination rule of evidence theory based on similarity of focal elements, Pattern Recognition and Artificial Intelligence 22 (2009) 169–175.
[42] T. Yang, Q.G. Li, Reduction about approximation spaces of covering generalized rough sets, International Journal of Approximation Reasoning 51 (2010) 335–345.
[43] Y.Q. Yao, J.S. Mi, Z.J. Li, Attribute reduction based o n generalized fuzzy evidence thoery in fuzzy decision system, Fuzzy Sets and Systems 170 (2011) 64–75.
[44] Y.Y. Yao, P.J. Lingras, Interpretations of belief functions in the theory of rough sets, Information Sciences 104 (1998) 81–106.

[45] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Information Sciences 111 (1998) 239–259.
[46] Y.Y. Yao, Constructive and algebraic methods of the theory of rough sets, Information Sciences 109 (1998) 21–47.
[47] L.A. Zadeh, Fuzzy sets, Information Control 8 (1965) 338–353.
[48] L.A. Zadeh, Probability measures of fuzzy events, Journal of Mathematical Analysis and Applications 23 (1968) 421–427.
[49] L.A. Zadeh, Fuzzy sets and information granularity, in: M. Gupta, R. Ragade, R.R. Yager (Eds.), Advances in Fuzzy Set Theory and Applications, North-Holland Publishing Co., Amsterdam, 1979, pp. 3–18.
[50] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1 (1978) 3–28.
[51] W. Zakowski, Approximation in the space (U, Π), Demonstratio Mathematica 16 (1983) 761–769.
[52] M. Zhang, L.D. Xu, W.X. Zhang, H.Z. Li, A rough set approach to knowledge reduction based on inclusion degree and evidence reasoning theory, Expert Systems 20 (2003) 298–304.
[53] W.X. Zhang, J.S. Mi, W.Z. Wu, Approaches to knowledge reductions in inconsistent systems, International Journal of Intelligent Systems 18 (2003) 989–1000.
[54] N. Zhong, Rough sets in knowledge discovery and data mining, Fuzzy Theory Systems 13 (2001) 581–591.
[55] P. Zhu, Covering rough sets based on neighborhoods: An approach without using neighborhoods, Journal International Journal of Approximate Reasoning 52 (2011) 461–472.
[56] W. Zhu, Generalized rough sets based on relations, Information Sciences 177 (2007) 4997–5011.
[57] W. Zhu, F.Y. Wang, Reduction and axiomization of covering generelized rough sets, Information Sciences 152 (2003) 217–230.
[58] W. Zhu, Relationship between generalized rough sets based on binary relation and covering, Information Sciences 179 (2009) 210–225.