International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems
Vol. , No. () 1–??

UNCERTAINTY MEASURES FOR MULTIGRANULATION APPROXIMATION SPACE

GUOPING LIN

Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, School of Computer and Information Technology, Shanxi University, Taiyuan, 030006, Shanxi, China
School of Mathematics and Statistics, Minnan Normal University, Minnan Normal University, Zhangzhou, 363000, Fujian, China

JIYE LIANG*, YUHUA QIAN

Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, School of Computer and Information Technology, Shanxi University, Taiyuan, 030006, Shanxi, China
∗ljy@sxu.edu.cn

Received (received date)
Revised (revised date)

Multigranulation rough set theory is a relatively new mathematical tool for solving complex problems in the multigranulation or distributed circumstances which are characterized by vagueness and uncertainty. In this paper, we first introduce the multigranulation approximation space. According to the idea of fusing uncertain, imprecise information, we then present three uncertainty measures: fusing information entropy, fusing rough entropy, and fusing knowledge granulation in the multigranulation approximation space. Furthermore, several essential properties (equivalence, maximum, minimum) are examined and the relationship between the fusion information entropy and the fusion rough entropy is also established. Finally, we prove these three measures are monotonously increasing as the partitions become finer. These results will be helpful for understanding the essence of uncertainty measures in multigranulation rough space and enriching multigranulation rough set theory.

Keywords: Rough set theory; multigranulation; fusion information entropy; fusion knowledge granulation; fusion rough entropy

1. Introduction

Rough set theory, proposed by Z. Pawlak[1,2], is a soft computing tool to deal with vagueness and uncertainty in many domains, including artificial intelligence and cognitive sciences. The rough set philosophy is based on the assumption that with every object of the universe there is associated a certain amount of information (data, knowledge), expressed by means of some attributes used for object description. Objects having the same description are indiscernible (similar) with respect to the available information by use of the indiscernibility relation. Classical definitions of lower and upper approximations were originally introduced with reference to the indiscernibility relation. Furthermore, the indiscernibility relation results in information granulation from which inconsistency or ambiguity is followed in the domain of discourse.

According to the fundamental importance of rough set theory in many applicable fields including artificial intelligence and cognitive sciences, many researchers have studied how to measure uncertainty in the rough set theory. The entropy of a system defined by Shannon[3] (1948) is a measure of uncertainty about the system's actual structure. It has been a useful mechanism for characterizing the information content in various modes and applications in many diverse fields. Since then, a lot of researchers (see, e.g.[4−9]) have used Shannon's entropy and its variants to measure uncertainty for rough set theory. However, Shannon's entropy cannot measure the fuzziness in rough set theory, therefore Liang[10,11] proposed a new information entropy. Unlike the logarithmic behavior of Shannon's entropy, the gain function of this entropy possesses the complement nature. It can be used to measure the fuzziness of rough set and rough classification. Mi et al. also gave a new fuzzy entropy and applied it for measuring the fuzziness of a fuzzy-rough set based partition (see, e.g.[12]). Qian and Liang presented combination information entropy with intuitionistic knowledge content characteristic in incomplete information systems and introduced a combination granulation to measure the uncertainty of an incomplete information system, and the relationship between these two concepts was established(see, e.g.[13,14]).

As a significant extension of Pawlak a single-granulation rough set, the multigranulation rough set (MGRS) was proposed by Qian et al. for a user's different requirements or targets of problem solving (see, e. g.[15]). One of important contributions in MGRS is to describe the lower and upper approximations of the rough set by multiple equivalence relations (multiple granulations) instead of a single equivalence relation (a single granulation). Since then, many researchers have extended the classical MGRS by using multiple generalized binary relations, such as neighborhood relation (see, e. g.[16]), tolerance relation (see, e. g.[17]), fuzzy relation(see, e. g.[18]) and variable precision relation (see, e. g.[19]). It is worthwhile to mention that Liang et al.(see, e. g.[20]) proposed an efficient rough feature selection algorithm for large-scale data sets by the view of multigranulation, which shows an important implication of MGRS.

However, the mechanism that how to characterize the vagueness and uncertainty for multigranulation approximation space has not been widely studied. Although many uncertainty measures for classical rough set theory were developed in (see, e.g.[4−11,18,21−26]), they are not applicable in some multigranulation or distributed circumstances. In this paper, by using the idea of fusing uncertain, imprecise information(see, e.g.[27−29]), we give a definition of fusing information entropy, fusing rough entropy, and fusing knowledge granulation for multigranulation approximation space and address some properties of them, respectively. Noting that these three measures are monotonic increasing as the partitions become finer, which is a essential property of uncertainty measure argued by Beaubouef et al(see, e.g.[6]). These results will be helpful for understanding the essence of uncertainty measure in the multigranulation rough space.

This paper aims to establish uncertainty measures for multigranulation rough set theory. Some preliminary concepts such as the knowledge base and the partial relation "$\leq$" are reviewed in Section 2. In Section 3, we give the definitions of the partial relation "$\leq^M$" and the multigranulation approximation space. Then we propose three measures: fusion information entropy, fusion knowledge granulation, fusion rough entropy. Some of their important properties are investigated and the relationship among fusion information entropy, fusion knowledge granulation, fusion rough entropy is established as well. In Section 4 concludes this paper with some remarks and discussions.

2. Preliminaries

In this section, we introduce some fundamental key concepts of rough set theory, the multigranulation rough set theory(see, e.g.[1,2,13]), and several uncertainty measures in the rough set theory(see, e.g.[1,2,13]). Throughout this paper, we suppose that the universe U is a finite nonempty set.

2.1. *Basic Concepts of Rough Set and Multigranulation Rough Set*

Let (U, AT) be an information system, where U is a non-empty finite set of objects, and AT is a non-empty finite set of attributes. Each subset of attributes $A \subseteq AT$ determines an equivalence relation R which constitutes a partition of U, denoted by $\widehat{A}$, or just U/R or U/A. We say the partition $\widehat{B}$ is coarser than the partition $\widehat{A}$(or the partition $\widehat{A}$ is finer than the partition $\widehat{B}$), denoted by $\widehat{A} \leq \widehat{B}$, if they satisfy the condition that

$$\forall A_i \in \widehat{A}, \exists B_j \in \widehat{B} \textit{ such that } A_i \subseteq B_j.$$

If $\widehat{A} \leq \widehat{B}$ and $\widehat{A} \neq \widehat{B}$, then we say $\widehat{B}$ is strictly coarser than $\widehat{A}$ (or $\widehat{A}$ is strictly finer than $\widehat{B}$) and written by $\widehat{A} < \widehat{B}$. In particular, in the studies about information systems, the identity partition $\omega(U)$ and universal partition $\delta(U)$ are given, where $\omega(U) = \{\{x\} \mid x \in U\}$ and $\delta(U) = \{U\}$. The former is the finest partition on the universe, and the latter is the roughest partition on the universe.

For any given information system (U, AT), where U is a non-empty finite set of objects, AT is a non-empty finite set of attributes. Suppose $A \subseteq AT$ and $X \subseteq U$, one can define a lower approximation and an upper approximation of X in U by

$$\underline{A}(X) = \bigcup\{A_i \in \widehat{A} \mid A_i \subseteq X\},$$

and

$$\overline{A}(X) = \bigcup\{A_i \in \widehat{A} \mid A_i \cap X \neq \emptyset\},$$

where $\underline{A}(X)$ is a set of objects that belong to X with certainty, while $\overline{A}(X)$ is a set of objects that possibly belong to X. It is easily shown that the lower approximation and the upper approximation of X in U are described by a single binary relation on the universe. In other words, this classical rough set theory is established by a single granulation in the view of granular computing(see, e. g.[30]). Here, we say $(U, \widehat{A})$ a single granulation space.

Recently, Qian et al. have extended Pawlak's rough set model to multigranulation rough set model(MGRS) according to a user's different requirements or targets of problem solving. In MGRS, the set approximations are defined by using multiple equivalence relations (multiple granulations) on the universe.

Let $S = (U, AT, f)$ be a complete information system, $X \subseteq U$ and $A_1, A_2, \cdots, A_m \subseteq AT$. A lower approximation and an upper approximation of X in U with respect to $A_1, A_2, \cdots, A_m$ are denoted by $\underline{\sum_{i=1}^m A_i}X$ and $\overline{\sum_{i=1}^m A_i}X$, respectively, where

$$\underline{\sum_{i=1}^{m}} A_iX = \bigcup\{x \in U \mid [x]_{A_i} \subseteq X, for\ some\ i \leq m\},$$

and

$$\overline{\sum_{i=1}^{m} A_i}X = \sim \underline{\sum_{i=1}^{m}} A_i(\sim X).$$

The lower approximation of a set X with respect to $\underline{\sum_{i=1}^m A_i}$ is a set of all elements, which can certainly be classified as X using $\underline{\sum_{i=1}^m A_i}$. The upper approximation of a set is the set of all elements can possibly be classical as X using $\underline{\sum_{i=1}^m A_i}$.

In [14], $(U, \widehat{A_1}, \widehat{A_2}, \cdots, \widehat{A_m})$ is called a knowledge base in which $\widehat{A_1}, \widehat{A_2}, \cdots, \widehat{A_m}$ represent various granulations, denoted by K_M, i.e., $K_M = \{\widehat{A_1}, \widehat{A_2}, \cdots, \widehat{A_m}\}$. Furthermore, we can say $\Omega = \{K_M \mid K_M = \{\widehat{A_1}, \widehat{A_2}, \cdots, \widehat{A_m}\}, A_i \in AT\}$ a multigranulation approximation space, which is a family set of knowledge based. In order to differentiate classical rough set and multigranulation rough set, we call K_M a multigranulation knowledge base.

Now we define a partial order in Ω. Let P_M, Q_M be two multigranulation knowledge base in Ω, and we define the multigranulation knowledge base Q_M is coarser than the multigranulation knowledge base P_M (or the multigranulation knowledge base P_M is finer than the multigranulation knowledge base Q_M), denoted by $P_M \preceq^M Q_M$, between two multigranulation knowledge bases satisfy the following condition:

$$P_M \preceq^M Q_M \Leftrightarrow \forall P_i \in P_M, \exists Q_j \in Q_M \ such \ that \ P_i \le Q_j,$$

where " $\le$ " is a partial relation between two partitions in pawlak rough set. If $P_M \preceq^M Q_M$ and $P_M \neq Q_M$, then we say Q_M is strictly coarser than P_M (or P_M is strictly finer than Q_M) and write $P_M <^M Q_M$. Moreover, if $\forall P_i \in P_M, \exists Q_j \in Q_M \ such \ that \ P_i = Q_j \ and \ \forall Q_i \in Q_M, \exists P_j \in P_M \ such \ that \ Q_i = P_j$, we call $P_M \approx Q_M$. Particularly, if $\widehat{A}_i = \omega(U)$, for all $i \in \{1, 2, \cdots, m\}$, we say $K_M = \{\omega(U), \omega(U), \cdots, \omega(U)\}$ is a finest multigranulation knowledge base in Ω, whereas, if for all $\widehat{A}_i = \delta(U)$, for all $i \in \{1, 2, \cdots, m\}$, we say $K_M = \{\delta(U), \delta(U), \cdots, \delta(U)\}$ is a coarsest multigranulation knowledge base in Ω.

2.2. *Information Entropy, Knowledge Granulation, and Rough entropy in the Rough Set Theory*

Definition 1.[5] (Information entropy) Let $K = (U, \widehat{A})$be a knowledge base, and $\widehat{A}$ a partition of U. An information entropy of knowledge $\widehat{A}$ in the rough set theory is defined by

$$IE(\widehat{A}) = -\sum_{i=1}^{m} \frac{|A_i|}{|U|} \log_2 \frac{|A_i|}{|U|}.$$

This granularity measure, *IE*, measures the uncertainty associated with the prediction of outcome where elements of each partition set $\widehat{A}$ are indistinguishable. $IE(\widehat{A})$ achieves the maximum value $\log_2 |U|$ if and only if $\widehat{A} = \omega(U)$. $IE(\widehat{A})$ achieves the minimum value 0 if and only if $\widehat{A} = \delta(U)$. Obviously, when $\widehat{A}$ is a partition of U, or an equivalence relation on U, we have that $0 \le IE(\widehat{A}) \le \log_2 |U|$.

Definition 2.[11] (Knowledge granulation) Let $K = (U, \widehat{A})$ be a knowledge base, and $\widehat{A}$ a partition of U. Granulation of knowledge $\widehat{A}$ in the rough set theory is defined by

$$GK(\widehat{A}) = \frac{1}{|U|^2} \sum_{i=1}^{m} |A_i|^2.$$

$GK(\widehat{A})$ achieves the minimum value $\frac{1}{|U|}$ if and only if $\widehat{A} = \omega(U)$. $GK(\widehat{A})$ achieves the maximum value 1 if and only if $\widehat{A} = \delta(U)$. Obviously, when $\widehat{A}$ is a partition of U, or an equivalence relation on U, we have that $\frac{1}{|U|} \leq GK(\widehat{A}) \leq 1$. Knowledge granulation can represent discernibility ability of knowledge, the smaller GK is, the stronger its discernibility ability.

Definition 3.[10] (Rough entropy) Let $K = (U, \widehat{A})$ be a knowledge base, and $\widehat{A}$ a partition of U. Rough entropy of knowledge $\widehat{A}$ in the rough set theory is defined by

$$E_r(\widehat{A}) = -\sum_{i=1}^{m} \frac{|A_i|}{|U|} \log_2 \frac{1}{|A_i|}.$$

Where $\frac{|A_i|}{|U|}$ represents the probability of equivalence class A_i within the universe U, $\frac{1}{|A_i|}$ denotes the probability of one of the values in equivalence class A_i. $E_r(\widehat{A})$ achieves the maximum value $\log_2 |U|$ if and only if $\widehat{A} = \omega(U)$. $E_r(\widehat{A})$ achieves the minimum value 0 if and only if $\widehat{A} = \delta(U)$. Obviously, when $\widehat{A}$ is a partition of U, or an equivalence relation on U, we have that $0 \leq E_r(\widehat{A}) \leq \log_2 |U|$.

3. Measures for Multigranulation Approximation Space

Multigranulation rough set theory is also a relatively new mathematical tool for use in computer applications in multigranulation or distributed circumstances which are characterized by vagueness and uncertainty. This approach seems to be of fundamental importance to artificial intelligence and cognitive sciences. However, the mechanism that how to characterize the vagueness and uncertainty for multigranulation approximation space has not been widely studied. It is necessary to investigate computational methods of uncertainty for multigranulation approximation space. In this section, we propose three uncertainty measures for it by utilizing the idea of uncertain information fusing method, accordingly, we call these measures fusing uncertainty measures.

Definition 4. (Fusion information entropy) Let $\Omega = \{K_M\}$ be a multigranulation approximation space, $P_M \in \Omega$ and $P_M = \{\widehat{A}_1, \widehat{A}_2, \cdots, \widehat{A}_m\}$, in which $\widehat{A}_i, i \in \{1, 2, \cdots, m\}$ is a partition of U. Then a fusion information entropy of the multigranulation knowledge base P_M is defined as follows:

$$\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = (\sum_{i=1}^{m} IE^2(\widehat{A}_i))^{\frac{1}{2}}.$$

We denote $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ by $\boldsymbol{MIE}(P_M)$. In the equation $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$, $IE(\widehat{A})$ is an information entropy of knowledge $\widehat{A}$. $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ is a Euclidean distance from the vector $\overrightarrow{P_M} = (IE(\widehat{A}_1), IE(\widehat{A}_2), \cdots, IE(\widehat{A}_m))$ to

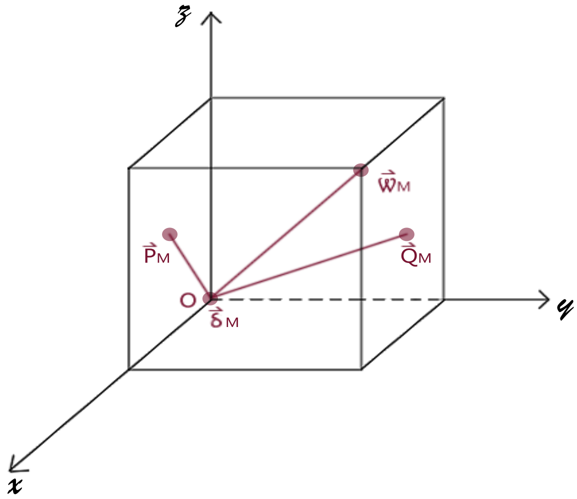

Fig. 1. The model of fusion information entropy

the vector $\overrightarrow{0} = (0, 0, \cdots, 0)$. The semantics of such a distance $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ is an uncertainty fusing measure which aims to represent an uncertainty extant from any a given multigranulation knowledge base to a finest base $P_M(\delta)$. The fusion information entropy represents the uncertainty associated with the prediction of outcomes with respect to m granulations. The vector $\overrightarrow{0}$ shown that a fusion information entropy of the multigranulation knowledge base is equal to 0 if and only if each single granulation is just the finest granulation $\delta(U)$.

To more easily understand fusion information entropy, we employ Figure 1 for further illustration. For simplicity, we suppose the multigranulation space constructed with three granular structures.

In Figure 1, coordinate of every point in the cube is regarded as a vector and its component coordinate represents information entropy for a granular structure. $\overrightarrow{\delta}_M$ means the vector $\overrightarrow{0} = \{0, 0, 0\}$ induced by the finest multigranulation approximation space which will be defined later. $\overrightarrow{\omega}_M$ means the vector induced by the coarsest multigranulation approximation space. $\overrightarrow{P}_M$ means the vector induced by some multigranulation space as well as $\overrightarrow{Q}_M$. The distance measure from a point $\overrightarrow{\delta}_M$ to a point $\overrightarrow{\omega}_M$(or $\overrightarrow{Q}_M$,$\overrightarrow{P}_M$) represents the uncertainty measure of the fusion information entropy of the multigranulation approximation space ω_M(or Q_M,P_M).

In particular, when $m = 1$, the fusion information entropy $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ will degenerate to the single granulation information entropy in the Pawlak rough set.

However, we wonder whether the fusion information entropy defined by the form of the square root of quadratic sum is more reasonable than that defined by the form of a simple linear combination of information entropy. In what follows, we employ an example to illustrate them.

For example, let $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ be a universe of discourse. Suppose $P_M = \{\widehat{A}_1, \widehat{A}_2\}$ and $Q_M = \{\widehat{B}_1, \widehat{B}_2\}$ are two multigranulation spaces, where $\widehat{A}_1 = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$, $\widehat{A}_2 = \{U\}$, $\widehat{B}_1 = \{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6, x_7, x_8\}\}$, and $\widehat{B}_2 = \{\{x_1, x_5, x_6, x_7\}, \{x_2, x_3, x_4, x_8\}\}$ four partitions of U. Then we obtain $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2) = \sqrt{IE^2(\widehat{B}_1) + IE^2(\widehat{B}_2)} = 2$ and $\boldsymbol{MIE}(\widehat{B}_1 \oplus \widehat{B}_2) = \sqrt{IE^2(\widehat{A}_1) + IE^2(\widehat{A}_2)} = 1$ according to Definition 1. However, if we compute them by using a simple linear combination of information entropy, i.e., $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2) = \frac{IE(\widehat{A}_1)+IE(\widehat{A}_1)}{2} = 1$ and $\boldsymbol{MIE}(\widehat{B}_1 \oplus \widehat{B}_2) = \frac{IE(\widehat{B}_1)+IE(\widehat{B}_1)}{2} = 1$. From the above result, one can find that two different multigranulation spaces have the same fusion information entropy it one employs a simple linear combination of information entropy. Whereas they are different by using Definition 1. These results have shown that the fusion entropy defined by the form of square root of a sum of squares is more meaningful than that defined by the form of a simple linear combination of information entropy.

Proposition 1. *For any* $i \in \{1, 2, \cdots, m\}$, *we have* $\mathrm{IE}(\widehat{A}_i) \leq \mathbf{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$.

Proof. Note that information entropy $IE(\widehat{A}_i) \geq 0$, we have $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = \sqrt{\sum_{i=1}^{m} IE^2(\widehat{A}_i)} \geq \sqrt{IE^2(\widehat{A}_i)} = IE(\widehat{A}_i)$. Hence, $IE(\widehat{A}_i) \leq \boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$.

Property 1. (Equivalence) Let $\Omega = \{K_M\}$ be a multigranulation approximation space, P_M, Q_M two multigranulation knowledge bases in Ω. If $P_M \approx Q_M$, then $\boldsymbol{MIE}(P_M) = \boldsymbol{MIE}(Q_M)$.

Proof. Let $P_M = \{\widehat{A}_1, \widehat{A}_2, \cdots, \widehat{A}_m\}$, $Q_M = \{\widehat{B}_1, \widehat{B}_2, \cdots, \widehat{B}_n\} \in \Omega$, suppose $P_M \approx Q_M$, we have $m = n$ and if for any $\widehat{A}_i \in P_M$, there exists $\widehat{B}_j \in Q_M$, we have $\widehat{A}_i = \widehat{B}_j$, then by the definition of partial relation " $\leq$ " and equivalence property of information entropy in Pawlak rough set[11], we have $IE(\widehat{A}_i) > IE(\widehat{B}_j)$ holds. Therefore, we have $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = (\sum_{i=1}^{m} IE^2(\widehat{A}_i))^{\frac{1}{2}} = (\sum_{i=1}^{m} IE^2(\widehat{B}_i))^{\frac{1}{2}} = \boldsymbol{MIE}(\widehat{B}_1 \oplus \widehat{B}_2 \oplus \cdots \oplus \widehat{B}_m)$. □

Property 2. (Maximum) Let $\Omega = \{K_M\}$ be a multigranulation approximation space, if each component value of the vector $\overrightarrow{K_M}$ is equal to $log_2|U|$ which means that the information entropy of each single granulation information entropy achieves the maximum value $log_2|U|$ in K_M, then $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the maximum value $\sqrt{m}log_2|U|$. On the other hand, $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the maximum value $\sqrt{m}log_2|U|$ if and only if $K_M = \{\omega(U), \omega(U), \cdots, \omega(U)\}$.

Property 3. (Minimum) Let $\Omega = \{K_M\}$ be a multigranulation approximation space, if each component value of the vector $\overrightarrow{K_M}$ is equal to 0, which means that the information entropy of each single granulation information entropy achieves the minimum value 0 in K_M, then $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the minimum value 0. On the other hand, $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the minimum value if and only if $K_M = \{\delta(U), \delta(U), \cdots, \delta(U)\}$. Obviously, when $\widehat{A}$ is a partition of U, or an equivalence relation on U, we have that $0 \leq IE(\widehat{A}) \leq \sqrt{m}\log_2|U|$.

Proposition 2. *Let $\Omega = \{K_M\}$ be a multigranulation approximation space, P_M, Q_M two knowledge bases in Ω, if $P_M <^M Q_M$, then* $\mathbf{MIE}(P_M) > \mathbf{MIE}(Q_M)$.

Proof. Suppose $P_M = \{\widehat{A}_1, \widehat{A}_2, \cdots, \widehat{A}_m\}$, $Q_M = \{\widehat{B}_1, \widehat{B}_2, \cdots, \widehat{B}_n\} \in \Omega$, since $P_M <^M Q_M$, we have $m > n$ and if for any $\widehat{A}_i \in P_M$, there exists $\widehat{B}_j \in Q_M$, we have $\widehat{A}_i < \widehat{B}_j$, by the monotonicity of information entropy in Pawlak rough set (see, e.g.[5]), $IE(\widehat{A}_i) > IE(\widehat{B}_j)$ holds. Therefore we have $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = (\sum_{i=1}^{m} IE^2(\widehat{A}_i))^{\frac{1}{2}} > (\sum_{i=1}^{n} IE^2(\widehat{B}_i))^{\frac{1}{2}} = \boldsymbol{MIE}(\widehat{B}_1 \oplus \widehat{B}_2 \oplus \cdots \oplus \widehat{B}_m)$. □

Proposition 2. states that fusing information entropy of multigranulation knowledge base monotonously increases as the partitions become finer.

Example 1. Let $U = \{x_1, x_2, x_3, x_4, x_5\}$ be a universe of discourse. Suppose $\widehat{A}_1 = \{\{x_1, x_2, x_3, x_4\}, \{x_5\}\}$ and $\widehat{A}_2 = \{\{x_1\}, \{x_2, x_4\}, \{x_3\}, \{x_5\}\}$ are two partitions of U. Then we compute the fusion information entropy of the multigranulation knowledge base $K_M = \{\widehat{A}_1, \widehat{A}_2\}$. By Definition 1, we have $IE(\widehat{A}_1) = \frac{4}{5}\log_2\frac{5}{4} + \frac{1}{5}\log_2 5$, $IE(\widehat{A}_2) = \frac{3}{5}\log_2 5 + \frac{2}{5}\log_2\frac{5}{2}$. Hence $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2) = \sqrt{IE^2(\widehat{A}_1) + IE^2(\widehat{A}_2)} = \sqrt{(\frac{4}{5}\log_2\frac{5}{4} + \frac{1}{5}\log_2 5)^2 + (\frac{3}{5}\log_2 5 + \frac{2}{5}\log_2\frac{5}{2})^2} = \frac{\sqrt{2}}{5}\sqrt{25\log_2^2 5 - 50\log_2 5 + 34}$.

Definition 5. (An axiom definition of the fusion knowledge granulation measures) Let $\Omega = \{K_M\}$ be a multigranulation approximation space, $\boldsymbol{MGK}$ be a mapping from the Ω to the set of non-negative real numbers $R^+ \cup 0$. We say that $\boldsymbol{MGK}$ is a fusion knowledge granulation of the multigranulation knowledge base in the multigranulation rough set if for any $P_M, Q_M \in \Omega$, and $\boldsymbol{MGK}$ satisfies the following conditions:

(1) $\boldsymbol{MGK}(P_M) \geq 0$ for any $P_M \in \Omega$; (Non-negativity)

(2) $\boldsymbol{MGK}(P_M) = \boldsymbol{MGK}(Q_M)$ for any $P_M, Q_M \in \Omega$ if there is $P_M \approx Q_M$; (Invariability)

(3) $\boldsymbol{MGK}(P_M) < \boldsymbol{MGK}(Q_M)$ for any $P_M, Q_M \in \Omega$ with $P_M <^M Q_M$. (Monotonicity)

Definition 6. (Fusion knowledge granulation) Let $\Omega = \{K_M\}$ be a multigranulation approximation space, $P_M \in \Omega$ and $P_M = \{\widehat{A}_1, \widehat{A}_2, \cdots, \widehat{A}_m\}$, in which $\widehat{A}_i, i \in \{1, 2, \cdots, m\}$ is a partition of U. Then a fusion knowledge granulation of the multigranulation knowledge base K_M is defined as follows:

$$\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = \frac{1}{m}(\sum_{i=1}^{m}(GK(\widehat{A}_i) - \frac{1}{|U|})^2)^{\frac{1}{2}}.$$

We denote $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ by $\boldsymbol{MGK}(P_M)$. In particular, when $m = 1$, the fusion information granulation $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \widehat{A}_m)$ will degenerate to the single information knowledge granulation in the Pawlak rough set.

Proposition 3. $\mathbf{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ *in Definition 6 is a knowledge granulation under Definition 5.*

Proof. (1) Obviously, $\boldsymbol{MGK}(P_M) \geq 0$, i.e., $\boldsymbol{MGK}(P_M) \geq 0$ is non-negative.

(2) Let $P_M = \{\widehat{A}_1, \widehat{A}_2, \cdots, \widehat{A}_m\}$, $Q_M = \{\widehat{B}_1, \widehat{B}_2, \cdots, \widehat{B}_n\} \in \Omega$, suppose $P_M \approx Q_M$, we have $m = n$ and if for any $\widehat{A}_i \in P_M$, there exists $\widehat{B}_j \in Q_M$, we have $\widehat{A}_i = \widehat{B}_j$, then by the definition of partial relation " $\leq$ " and equivalence property of knowledge granulation in Pawlak rough set[11], we have $GK(\widehat{A}_i) > GK(\widehat{B}_j)$ holds. Therefore, we have $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = \frac{1}{m}(\sum_{i=1}^{m}(GK(\widehat{A}_i) - \frac{1}{|U|})^2)^{\frac{1}{2}} = \frac{1}{m}(\sum_{i=1}^{m}(GK(\widehat{B}_i) - \frac{1}{|U|})^2)^{\frac{1}{2}}. = \boldsymbol{MGK}(\widehat{B}_1 \oplus \widehat{B}_2 \oplus \cdots \oplus \widehat{B}_m)$.

(3) Let $P_M, Q_M \in \Omega$, $P_M = \{\widehat{A}_1, \widehat{A}_2, \cdots, \widehat{A}_m\}$, $Q_M = \{\widehat{B}_1, \widehat{B}_2, \cdots, \widehat{B}_n\} \in \Omega$ with $P_M <^M Q_M$, If $P_M <^M Q_M$, we have $m > n$ and if for any $\widehat{A}_i \in P_M$, there exists $\widehat{B}_j \in Q_M$, we have $\widehat{A}_i < \widehat{B}_j$, by the monotonicity of knowledge granulation in Pawlak rough set[11], $GK(\widehat{A}_i) > GK(\widehat{B}_j)$ holds. Therefore we have $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = \frac{1}{m}(\sum_{i=1}^{m}(GK(\widehat{A}_i) - \frac{1}{|U|})^2)^{\frac{1}{2}} < \frac{1}{m}(\sum_{i=1}^{m}(GK(\widehat{B}_i) - \frac{1}{|U|})^2)^{\frac{1}{2}} = \boldsymbol{MGK}(\widehat{B}_1 \oplus \widehat{B}_2 \oplus \cdots \oplus \widehat{B}_m)$. □

Proposition 4. *For any* $i \in \{1, 2, \cdots, m\}$*, we have* $\mathrm{GK}(\widehat{A}_i) \leq \mathbf{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$.

Proof. The proof follows from Definition 6. □

Property 4. (Maximum) Let $\Omega = \{K_M\}$ be a multigranulation approximation space, if each component value of the vector $\overrightarrow{K_M}$ is equal to 1 , which means that each single information knowledge granulation achieves the maximum value 1 in K_M, then $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \widehat{A}_m)$ achieves the maximum value $\frac{\sqrt{m}}{m}(1 - \frac{1}{|U|})$. On the other hand, $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the minimum value if and only if $K_M = K_M(\delta) = \{\delta(U), \delta(U), \cdots, \delta(U)\}$.

Property 5. (Minimum) Let $\Omega = \{K_M\}$ be a multigranulation approximation space, if each component value of the vector $\overrightarrow{K_M}$ is equal to 0, which means that each single information knowledge granulation achieves the minimum value $\frac{1}{|U|}$ in K_M, then $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \widehat{A}_m)$ achieves the minimum value 0. On the other hand, $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the minimum value if and only if $K_M = K_M(\omega) = \{\omega(U), \omega(U), \cdots, \omega(U)\}$. Obviously, when $\widehat{A}$ is a partition of U, or an equivalence relation on U, we have that $0 \leq GK(\widehat{A}) \leq \frac{\sqrt{m}}{m}(1 - \frac{1}{|U|})$.

Example 2. (Continued from Example 1) Let $U = \{x_1, x_2, x_3, x_4, x_5\}$, $\widehat{A}_1 = \{\{x_1, x_2, x_3, x_4\}, \{x_5\}\}$ and $\widehat{A}_2 = \{\{x_1\}, \{x_2, x_4\}, \{x_3\}, \{x_5\}\}$ two partitions of U, then we compute the fusion knowledge granulation of the multigranulation knowledge base $K_M = \{\widehat{A}_1, \widehat{A}_2\}$. By Definition 2, we have $GK(\widehat{A}_1) = \frac{17}{25}$, $GK(\widehat{A}_2) = \frac{7}{25}$. Hence $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2) = \sqrt{GK^2(\widehat{A}_1) + GK^2(\widehat{A}_2)} = \frac{3}{25}\sqrt{5}$.

Definition 7. (Fusion rough entropy) Let $\Omega = \{K_M\}$ be a multigranulation approximation space, $P_M \in \Omega$ and $P_M = \{\widehat{A}_1, \widehat{A}_2, \cdots, \widehat{A}_m\}$, in which $\widehat{A}_i, i \in \{1, 2, \cdots, m\}$ is a partition of U. Then a fusion information entropy of the multigranulation knowledge base K_M is defined as follows:

$$\boldsymbol{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \widehat{A}_m) = (\sum_{i=1}^{m} E_r^2(\widehat{A}_i))^{\frac{1}{2}}.$$

We denote $\boldsymbol{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \widehat{A}_m)$ by $\boldsymbol{ME}_r(P_M)$. In the equation of $\boldsymbol{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \widehat{A}_m)$, $IE(\widehat{A})$ is an information entropy of knowledge $\widehat{A}$. $\boldsymbol{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \widehat{A}_m)$ represents a Euclidean distance from the vector $\overrightarrow{K_M} = (E_r(\widehat{A}_1), E_r(\widehat{A}_2), \cdots E_r(\widehat{A}_m))$ to the vector $\overrightarrow{0} = (0, 0, \cdots, 0)$. The vector $\overrightarrow{0}$ denotes a fusion information entropy of the multigranulation knowledge base is equal to 0 if and only if each single granulation information entropy is equal to 0.

In particular, when $m = 1$, the fusion information entropy $\boldsymbol{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ will degenerate to the single granulation information entropy in the Pawlak rough set.

Proposition 5. *For any $i \in \{1, 2, \cdots, m\}$, we have* $\mathrm{E}_r(\widehat{A}_i) \leq \mathbf{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$.

Proof. The proof follows from Definition 7. □

Property 6. (Equivalence) *Let $\Omega = \{K_M\}$ be a multigranulation approximation space, P_M, Q_M two knowledge bases in Ω, if $P_M \approx Q_M$, then* $\mathbf{ME}_r(P_M) = \mathbf{ME}_r(Q_M)$.

Proof. The proof follows from Property 1.

Property 7. (Maximum) *Let $\Omega = \{K_M\}$ be a multigranulation approximation space, if each component value of the vector $\overrightarrow{K_M}$ is equal to $log_2|U|$, which means that the rough entropy of each single information granulation achieves the maximum value $log_2|U|$ in K_M, then $\mathbf{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the maximum value $\sqrt{m}log_2|U|$. On the other hand, $\mathbf{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the maximum value if and only if $K_M = \{\omega(U), \omega(U), \cdots, \omega(U)\}$.*

Property 8. (Minimum) *Let $\Omega = \{K_M\}$ be a multigranulation approximation space, if each component value of the vector $\overrightarrow{K_M}$ is equal to 0, which means that the rough entropy of each single information granulation achieves the minimum value 0 in K_M, then $\mathbf{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the minimum value 0. On the other hand, $\mathbf{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)$ achieves the minimum value if and only if $K_M = \{\delta(U), \delta(U), \cdots, \delta(U)\}$.*

Obviously, when $\widehat{A}$ is a partition of U, or an equivalence relation on U, we have that $0 \leq \boldsymbol{ME}_r(\widehat{A}) \leq \sqrt{m}log_2|U|$.

Proposition 6. *Let $\Omega = \{K_M\}$ be a multigranulation approximation space, P_M, Q_M two knowledge bases in Ω, if $P_M <^M Q_M$, then $\mathbf{ME}_r(P_M) < \mathbf{ME}_r(Q_M)$.*

Proof. Let $P_M = \{\widehat{A}_1, \widehat{A}_2, \cdots, \widehat{A}_m\}$, $Q_M = \{\widehat{B}_1, \widehat{B}_2, \cdots, \widehat{B}_n\} \in \Omega$, since $P_M <^M Q_M$, we have $m > n$ and if for any $\widehat{A}_i \in P_M$, there exists $\widehat{B}_j \in Q_M$, we have $\widehat{A}_i < \widehat{B}_j$, by the monotonicity of information entropy in Pawlak rough set [11], $IE(\widehat{A}_i) > IE(\widehat{B}_j)$ holds. Therefore we have $\boldsymbol{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = (\sum_{i=1}^{m} IE_r^2(\widehat{A}_i))^{\frac{1}{2}} > (\sum_{i=1}^{n} IE_r^2(\widehat{B}_i))^{\frac{1}{2}} = \boldsymbol{ME}_r(\widehat{B}_1 \oplus \widehat{B}_2 \oplus \cdots \oplus \widehat{B}_m)$. □

Proposition 6 states that fusing information entropy of multigranulation knowledge base monotonously increases as the partitions become finer.

Proposition 7. *Let $\Omega = \{K_M\}$ be a multigranulation approximation space, $P_M \in \Omega$ and $P_M = \{\widehat{A}_1, \widehat{A}_2, \cdots, \widehat{A}_m\}$, in which $\widehat{A}_i, i \in \{1, 2, \cdots, m\}$ is a partition of U. Then the relation between the fusion information entropy and the fusion rough entropy of the multigranulation knowledge base P_M is as follows:*

$$\mathbf{MIE}^2(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) + 2\sum_{i=1}^{m} \mathrm{IE}(\widehat{A}_i)\mathrm{E}_r(\widehat{A}_i) + \mathbf{ME}_r^2(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = m \log_2^2 |U|.$$

Proof.

$$
\begin{aligned}
&\boldsymbol{MIE}^2(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) + \boldsymbol{ME}_r^2(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m)\\
&= \sum_{i=1}^m IE^2(\widehat{A}_i) + \sum_{i=1}^m E_r^2(\widehat{A}_i)\\
&= \sum_{i=1}^m [(IE(\widehat{A}_i) + E_r(\widehat{A}_i))^2 - 2IE(\widehat{A}_i) \bullet E_r(\widehat{A}_i)]\\
&= \sum_{i=1}^m (-\sum_{i=1}^m \tfrac{|A_i|}{|U|} \log_2 \tfrac{|A_i|}{|U|} + E_r(\widehat{A}_i))^2 - 2IE(\widehat{A}_i) \bullet E_r(\widehat{A}_i))\\
&= \sum_{i=1}^m (-\sum_{i=1}^m \tfrac{|A_i|}{|U|} (\log_2 |A_i| - \log_2 |U|) + E_r(\widehat{A}_i))^2 - 2IE(\widehat{A}_i) \bullet E_r(\widehat{A}_i))\\
&= \sum_{i=1}^m (-\sum_{i=1}^m \tfrac{|A_i|}{|U|} (-\log_2 |A_i| + \log_2 \tfrac{1}{|U|}) + E_r(\widehat{A}_i))^2 - 2IE(\widehat{A}_i) \bullet E_r(\widehat{A}_i))\\
&= \sum_{i=1}^m (-\sum_{i=1}^m \tfrac{|A_i|}{|U|} (-1) \log_2 |A_i| + \sum_{i=1}^m \tfrac{|A_i|}{|U|} \log_2 \tfrac{1}{|U|}) + E_r(\widehat{A}_i))^2 - 2IE(\widehat{A}_i) \bullet E_r(\widehat{A}_i))\\
&= \sum_{i=1}^m (-E_r(\widehat{A}_i) + \sum_{i=1}^m \tfrac{|A_i|}{|U|} \log_2 \tfrac{1}{|U|}) + E_r(\widehat{A}_i))^2 - 2IE(\widehat{A}_i) \bullet E_r(\widehat{A}_i))\\
&= m \log_2^2 |U| - 2 \sum_{i=1}^m (IE(\widehat{A}_i) \bullet E_r(\widehat{A}_i)).
\end{aligned}
$$

Therefore, $\boldsymbol{MIE}^2(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) + 2\sum_{i=1}^m IE(\widehat{A}_i)E_r(\widehat{A}_i) + \boldsymbol{ME}_r^2(\widehat{A}_1 \oplus \widehat{A}_2 \oplus \cdots \oplus \widehat{A}_m) = m \log_2^2 |U|$. □

Example 3. Let $U = \{x_1, x_2, x_3, x_4, x_5\}$ be a universe of discourse. Suppose $\widehat{A}_1 = \{\{x_1, x_2, x_3, x_4\}, \{x_5\}\}$ and $\widehat{A}_2 = \{\{x_1\}, \{x_2, x_4\}, \{x_3\}, \{x_5\}\}$ are two partitions of U. Then we compute the fusion information entropy of the multigranulation knowledge base $K_M = \{\widehat{A}_1, \widehat{A}_2\}$. By Definition 7, we have $E_r(\widehat{A}_1) = \frac{4}{5}\log_2 \frac{5}{4}$, $E_r(\widehat{A}_2) = \frac{2}{5}\log_2 2$. Hence $\boldsymbol{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2) = \sqrt{E_r^2(\widehat{A}_1) + E_r^2(\widehat{A}_2)} = \frac{2\sqrt{7}}{5}\log_2 2$.

Example 4. Let $U = \{x_1, x_2, x_3, x_4, x_5\}$ be a universe of discourse. Suppose $P_M = \{\widehat{A}_1, \widehat{A}_2\}$, in which $\widehat{A}_1 = \{\{x_1, x_2\}, \{x_3\}, \{x_4, x_5\}\}$ and $\widehat{A}_2 = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4, x_5\}\}$, $Q_M = \{\widehat{B}_1, \widehat{B}_2\}$ in which $\widehat{B}_1 = \{\{x_1, x_2, x_3\}, \{x_4, x_5\}\}$ and $\widehat{B}_2 = \{\{x_1, x_2\}, \{x_3, x_4, x_5\}\}$. They are four partitions of U. Then we compute

(1) $IE(\widehat{A}_1) = \frac{1}{5}(5\log_2 5 - 4)$, $IE(\widehat{A}_2) = \frac{1}{5}(5\log_2 5 - 2)$, we have $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2) = \sqrt{IE(\widehat{A}_1) + IE(\widehat{A}_2)} = \frac{1}{5}(50\log_2^2 5 - 60\log_2 5 + 20)$. And $IE(\widehat{B}_1) = \frac{1}{5}(5\log_2 5 - 3\log_2 3 - 2)$, $IE(\widehat{B}_2) = \frac{1}{5}(5\log_2 5 - 3\log_2 3 - 2)$, we have that $\boldsymbol{MIE}(\widehat{B}_1 \oplus \widehat{B}_2) = \sqrt{GK^2(\widehat{B}_1) + GK^2(\widehat{B}_2)} = \frac{\sqrt{2}}{5}(5\log_2 5 - 3\log_2 3 - 2)$. Therefore, $\boldsymbol{MIE}^2(\widehat{A}_1 \oplus \widehat{A}_2) - \boldsymbol{MIE}^2(\widehat{B}_1 \oplus \widehat{B}_2) = \frac{2}{5}(4\log_2 5 - 3) > 0$. Note that $\boldsymbol{MIE}(\widehat{A}_i) \geq 0$, then $\boldsymbol{MIE}(\widehat{A}_1 \oplus \widehat{A}_2) > \boldsymbol{MIE}(\widehat{B}_1 \oplus \widehat{B}_2)$.

(2) $GK(\widehat{A}_1) = \frac{9}{25}$, $GK(\widehat{A}_2) = \frac{7}{25}$, we have $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2) = \sqrt{GK^2(\widehat{A}_1) + GK^2(\widehat{A}_2)} = \frac{\sqrt{5}}{5}$. And $GK(\widehat{B}_1) = \frac{13}{25}$, $GK(\widehat{B}_2) = \frac{13}{25}$, we have $\boldsymbol{MGK}(\widehat{B}_1 \oplus \widehat{B}_2) = \sqrt{GK^2(\widehat{B}_1) + GK^2(\widehat{B}_2)} = \frac{4\sqrt{2}}{5}$. Therefore, $\boldsymbol{MGK}^2(\widehat{A}_1 \oplus \widehat{A}_2) - \boldsymbol{MGK}^2(\widehat{B}_1 \oplus \widehat{B}_2) = -\frac{27}{25} < 0$. Note that $\boldsymbol{MGK}(\widehat{A}_i) \geq 0$, then $\boldsymbol{MGK}(\widehat{A}_1 \oplus \widehat{A}_2) < \boldsymbol{MGK}(\widehat{B}_1 \oplus \widehat{B}_2)$.

(3) $E_r(\widehat{A}_1) = \frac{4}{5}$, $E_r(\widehat{A}_2) = \frac{2}{5}$, we have $\boldsymbol{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2) = \sqrt{E_r^2(\widehat{A}_1) + E_r^2(\widehat{A}_2)} = \frac{2\sqrt{5}}{5}$. And $E_r(\widehat{B}_1) = \frac{1}{5}(3\log_2 3 + 2)$, $E_r(\widehat{B}_2) = \frac{1}{5}(2 + 3\log_2 3)$. We have $\boldsymbol{ME}_r(\widehat{B}_1 \oplus$

$\widehat{B}_2) = \sqrt{E_r^2(\widehat{B}_1) + E_r^2(\widehat{B}_2)} = \frac{\sqrt{2}}{5}(3\log_2 3 + 2)$. Therefore, $\boldsymbol{ME}_r^2(\widehat{A}_1 \oplus \widehat{A}_2) - \boldsymbol{ME}_r^2(\widehat{B}_1 \oplus \widehat{B}_2) = -\frac{2}{25}(9\log_2^2 3 - 12\log_2 3) < 0$. Note that $\boldsymbol{ME}_r(\widehat{A}_i) \geq 0$, then $\boldsymbol{ME}_r(\widehat{A}_1 \oplus \widehat{A}_2) < \boldsymbol{ME}_r(\widehat{B}_1 \oplus \widehat{B}_2)$.

4. Conclusions and Discussions

Multigranulation rough set theory is an importance extension of Pawlak rough set theory. This theory seems to be of fundamental importance to analyse rough data and complex problem solving in artificial intelligence and cognitive sciences domain. It is necessary to investigate to measure uncertainty in multigranulation rough set. In this paper, using the idea of information fusion, we have first addressed an investigation of uncertainty measure method for the multigranulation approximation space. More specifically, we have presented three uncertainty measures such as fusing information entropy, fusing rough entropy and fusing knowledge granulation. In addition, we have established their some nice properties (equivalence, maximum, minimum) and proved that these three measures are monotonic increasing as the partitions become finer. These results will be helpful for understanding the essence of uncertainty measure in the multigranulation rough space and enriching MGRS theory.

Acknowledgements

The authors would like to thank the anonymous reviewers and the editor for their constructive and valuable comments. This work is supported by grants from National Natural Science Foundation of China under Grant (Nos. 71031006, 61322211, 61303131, 61379021), National Key Basic Research and Development Program of China (973) (No. 2007CB311002), the Construction Project of the Science and Technology Basic Condition Platform of Shanxi(No.2012091002-0101), and Education Committee of Fujian Province under Grant (Nos. JA13204, 2013J01028,2013J01029,ZZ2013j04).

References

1. Z. Pawlak, “Rough sets”, *Int. J. comput. Inform. Sci.* **11** (1982) 341-365.
2. Z. Pawlak, “Rough Sets: Theoretical aspects of reasoning about data”, *Kluwer Academic Publishers, Dordecht* 1991.
3. C. E. Shannon, “The mathematical theory of communication”, *The Bell System Technical Journal* **27** (3 and 4) (1948) 373-423 (see also 623-656).
4. I. Düntsch and G. Gediga, “Uncertainty measures of rough set prediction” , *Artif. Intell.* **106** (1998) 109-137.
5. M. J. Wierman, “Measuring uncertainty in rough set theory”, *Int. J. Gen. Syst.* **28** (1999) 283-297.
6. T. Beaubouef, F. E. Petry and G. Arora, “Information-theoretic measures of uncertainty for rough sets and rough relational databases”, *Inform. Sci.* **109** (1998) 185-195.

7. J. Klir and M. J. Wierman, *Uncertainty-based information: elements of generalished information theory*, Physica-Verlag, New York, 1999.
8. J. Y and K. S. Qu, "Information measures of roughness of knowledge and rough sets in incomplete information systems", *J. Syst. Sci. Syst. Engineer.* **24**(5) (2001) 544-547.
9. J. Y. Liang and Z. B. Xu, "The algorithm on knowledge reduction in incomplete information systems", *Int. J. Uncertai., Fuzz. Knowledge-Based Syst.* **24**(1) (2002) 95-103.
10. J. Y. Liang and Z. Z. Shi, "The information entropy, rough entropy and knowledge granulation in rough set theory", *Int. J. Uncertain., Fuzz. Knowledge-Based Syst.* **12**(1) (2004) 37-46.
11. J. Y. Liang, Shi. Z. Z and D. Y. Li, "The information entropy, rough entropy and knowledge granulation in incomplete systems", *Int. J. Gen. Syst.* **35**(6) (2006) 641-654.
12. J. S. Mi, Y. Leung and W. Z. Wu, "An uncertainty measure in partition-based sets", *Int. J. Gen. Syst.* **34**(1) (2005) 77-90.
13. Y. H. Qian and J. Y. Liang, "Combination entropy and combination granulation in incomplete information system", *Lecture Notes in Artifi. Intell.* **4062** (2006) 184-190
14. Y. H. Qian and J. Y. Liang, "Knowledge structure, knowledge granulation and knowledge distance in a knowledge base", *Int. J. Approx. Reason.*, **50** (2009) 174-188.
15. Y. H. Qian, J. Y. Liang, Y. Y. Yao and C. Y. Dang, "MGRS: A multi-granulation rough set", *Inform. Sci.* **180** (2010) 949-970.
16. G. P. Lin, Y. H. Qian and J. J. Li, "NMGRS: Neighborhood-based multigranulation rough sets", *Int. J. Approx. Reason.* **53**(7) (2012) 1080-1093.
17. Y. H. Qian, J. Y. Liang, Y. Y. Yao and C. Y. Dang, "Incomplete mutigranulation rough set", *IEEE T. Syst., Man, and Cybernet., Part A* **20** (2010) 420-430.
18. X. B. Yang, X. N. Song, H. L. Dou and J. Y. Yang, "Multi-granulation rough set: from crisp to fuzzy case", *Annals of Fuzzy Mathematics and Informatics* **1** (1) (2011) 55-70
19. W. H. Xu, Q. R. Wang and X. T. Zhang, "Multi-granulation fuzzy rough sets in a fuzzy tolerance approximation space", *Int. J. Fuzz. Syst.* **13**(4) (2011) 246-259.
20. J. Y. Liang, F. Wang, C. Y. Dang and Y. H. Qian, "An efficient rough feature selection algorithm with a multi-granulation view", *Int. J. Approx. Reason.* **53**(7) (2012) 1080-1093.
21. J. Y. Liang and D. Y. Li, " Uncertainty and knowledge acquisition in information system", Science Press, Beijing, 2005 (in Chinese).
22. Y. H. Qian and J. Y. Liang, "A new method for measuring the uncertainty in incomplete information systems,", *Int. J. Uncertain. Fuzz. Knowledge-Based Syst.* **4062** (2006) 184-190.
23. J. H. Wang, J. Y. Liang, Y. H. Qian and C. Y. Dang, "Uncertainty measure of rough sets based on a knowledge granulation for incomplete information system",*Int. J. Uncertain. Fuzz. Knowledge-Based Syst.* **17**(16) (2009) 855-880.
24. D.Q. Miao and J. Wang, "On the relationships between information entropy and roughness of knowledge in rough set theory", *Pattern Recog. Artif. Intell.* **11** (1998) 34-40 (in chinese).
25. T.T. Lee, "An information-theoretic analysis of relational databases. Part I: data dependencies and information metric", *IEEE T. Software Engineee.* **13** (1987) 1049-1061.
26. Y.Y Yao and L. Zhao "A measurement theory view on the granularity of partitions", *Inform. Sci.* **213**(2012) 1-13.
27. S. Maskell, " Bayesian approach to fusing uncertain, imprecise and conflicting infor-

mation", *Inform. Fusion* **9** (2008) 259-277.
28. D. R. Yu, Q. H. Hu and C. X. Wu, "Short communication uncertainty measures for fuzzy relations and their applications", *Appl. Soft Comput.* **7** (2007) 1135-1143.
29. K. Sommer, O. Khn, F. Puente Len and B. Siebert, "A Bayesian approach to information fusion for evaluating the measurement uncertainty", *Robot. Auton. Syst.* (2009) 339-344.
30. Y. Y. Yao, "Information granulation and rough approximation", *Int. J. Intell. Syst.* **16** (2001) 87-104.