Yang XB, Qian YH, Yang JY. Hierarchical structures on multigranulation spaces. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 27(6): 1169–1183 Nov. 2012. DOI 10.1007/s11390-012-1294-0

Hierarchical Structures on Multigranulation Spaces

Xi-Bei Yang[1,2] (杨习贝), Yu-Hua Qian[3,4] (钱宇华), *Member, CCF, IEEE*,
and Jing-Yu Yang[2] (杨静宇), *Member, CCF, IEEE*

[1] *School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China*

[2] *School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China*

[3] *Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education Taiyuan 030006, China*

[4] *School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China*

E-mail: yangxibei@hotmail.com; jinchengqyh@126.com; yangjy@mail.njust.edu.cn

Received September 13, 2011; revised April 6, 2012.

Abstract Though many hierarchical structures have been proposed to analyze the finer or coarser relationships between two granulation spaces, these structures can only be used to compare the single granulation spaces. However, it should be noticed that the concept of multigranulation plays a fundamental role in the development of granular computing. Therefore, the comparison between two multigranulation spaces has become a necessity. To solve such problem, two types of the multigranulation spaces are considered: one is the partition-based multigranulation space, the other is the covering-based multigranulation space. Three different hierarchical structures are then proposed on such two multigranulation spaces, respectively. Not only the properties about these hierarchical structures are discussed, but also the relationships between these hierarchical structures and the multigranulation rough sets are deeply investigated. It is shown that the first hierarchical structure is consistent with the monotonic varieties of optimistic multigranulation rough set, and the second hierarchical structure is consistent to the monotonic varieties of pessimistic multigranulation rough set, the third hierarchical structure is consistent to the monotonic varieties of both optimistic and pessimistic multigranulation rough sets.

Keywords hierarchical structure, multicovering rough set, multigranulation rough set, multigranulation space

1 Introduction

Though Granular Computing (GrC) was firstly proposed by Zadeh in fall 1996, the basic idea of granular computing had been mentioned by John Von Neumann, in 1951. In his "The General and Logical Theory of Automata"[1], it is noted that "The natural systems are of enormous complexity, and it is clearly necessary to subdivide the problem that they represent into several parts." In granular computing theory, an information granule is a clump of objects drawn together by indistinguishability, similarity, and proximity of functionality[2-4]. A granulation structure is a collection of the available information granules for problem solving (granulation structure is also referred as the granulation space in this paper). A quotient structure[5-6] is the abstract of granulation structure, in which not only each information granule is regarded as a point, but also the relationships among information granules in granulation space are transformed into the relationships among points. A knowledge structure is an improvement of quotient structure, in which each point is replaced by a meaningful name. A linguistic structure is not a mathematical formalism but a natural language formulation, and it is an unexplored field presently.

Granular computing has three different views: knowledge engineering, uncertainty theory, and how to solve/compute[5]. Though granular computing is motivated by uncertainty, it has been widely explored through knowledge engineering view, i.e., the information granules are regarded as the basic vehicles of knowledge.

Presently, although a well–accepted definition of granular computing has not been presented, researchers have obtained many gratifying results. These results show that the current research is mainly dominated by rough sets[7], fuzzy sets and quotient space theory[8].

Regular Paper

This work is supported by the National Natural Science Foundation of China under Grant Nos. 61100116, 61103133, the Natural Science Foundation of Jiangsu Province of China under Grant No. BK2011492, the Natural Science Foundation of Jiangsu Higher Education Institutions of China under Grant No. 11KJB520004, the Postdoctoral Science Foundation of China under Grant No. 20100481149, and the Postdoctoral Science Foundation of Jiangsu Province of China under Grant No. 1101137C.

It is worth mentioning that rough set has been demonstrated to be useful in data mining, knowledge discovering, pattern recognition, medical diagnose and so on. In Pawlak's rough set theory, the information granules are equivalence classes, the granulation space is the partition on the universe of discourse, the quotient structure is the abstract of such partition, in which any two points (abstract of any two equivalence classes) are disjoint, and the knowledge structure is the domain of the attribute in Pawlak's knowledge representation system[7] (information system) since each object holds one and only one value (name) on such attribute.

Obviously, rough set and most of the expanded rough sets are constructed through one and only one binary relation or covering. For such reason, we may call these rough sets the single granulation rough sets. In single granulation rough sets, a partition is a granulation space, a binary neighborhood system[9] induced by a binary relation is a granulation space, and a covering[10-12] is also a granulation space. Nevertheless, this is not completely consistent to the basic thinking of granular computing since granular computing emphasizes the granulation space characterized by multilevel and multiview, i.e. we often need to describe concurrently a target concept from some independent environments[13]. To solve such problem, Qian *et al.*[14-16] proposed the concept of the multigranulation rough set. The main difference between single granulation and multigranulation rough sets is that we can use multi-different sets of information granules for the approximating of concept. Since each set of information granules can be considered as a granulation space, the space induced by multi-different sets of information granules are referred as the multigranulation space.

Presently, the multigranulation approach progresses rapidly. For example, in Qian *et al.*'s multigranulation rough set theory, there are two basic models: one is the optimistic multigranulation rough set, the other is the pessimistic multigranulation rough set[16]. Following Qian *et al.*'s work, Yang *et al.* generalized multigranulation rough sets into fuzzy environment and then proposed the multigranulation fuzzy rough sets in [17]. Wu and Leung[18] investigated the multi–scale information system, which reflects the explanation of the same problem at different scales (levels of granulations). In [19], Qian *et al.* also proposed a positive approximation, which can be used to accelerate a heuristic process of attribute reduction. Since the positive approximation uses a preference ordering, which can make the granulation space finer step by step, positive approximation also reflects the thinking of multigranulation. Moreover, Tsau Young Lin, proposed the neighborhood system based model for granular computing and the corresponding rough set[6]. Generally speaking, a neighborhood system contains a set of neighborhoods, if each neighborhood is considered as an information granule, then the neighborhood system is also a useful model in multigranulation approach. From discussions above, we can see that multigranulation approach is a powerful tool in granular computing.

The purpose of this paper is to study the hierarchial structures on multigranulation spaces. This is mainly because most of hierarchical structures are proposed to deal with single granulation spaces. For example, Yao[20] suggested the use of hierarchical granulations for the study of stratified rough set approximations; Wang *et al.*[21] proposed a hierarchical knowledge space chain in terms of different knowledge granulation levels; Huang *et al.* and Zhang *et al.* proposed hierarchical structures on covering-based granulation space in [22] and [23], respectively. Zhang *et al.*[24] constructed a hierarchical structure on fuzzy quotient spaces through constructing normalized isosceles distance function and then proposed the uncertainty measurement of hierarchical quotient spaces in [25]. Liang *et al.*[26-27] analyzed the varieties of information entropy, rough entropy and knowledge granulation in terms of the variety of levels in hierarchical structure. Qian *et al.*[28-29] defined some operations on granulation spaces and then analyzed the finer or coarser relationships between original and operated granulation spaces. Different from the hierarchical structures we mentioned above, what will be discussed in this paper is based on the comparisons between two different multigranulation spaces. We may desire to find hierarchical structures on multigranulation spaces in terms of the varieties of multigranulation rough sets. We will also address such topic from the viewpoints of partition- and covering-based multigranulation rough sets, respectively.

To facilitate our discussions, we first present basic notions of partition-based rough set and multigranulation rough sets in Section 2. In Section 3, three different hierarchical structures are defined on partition-based multigranulation spaces (PBMS). Not only the relationships among these hierarchical structures are discussed, but also the relationships between these hierarchical structures and multigranulation rough sets are investigated. In Section 4, the optimistic and pessimistic rough sets are introduced into covering-based multigranulation space (CBMS). Similar to the partition case, in Section 5, three different hierarchical structures are also defined and analyzed on covering-based multigranulation spaces. Results are summarized in Section 6.

2 Preliminary Knowledge on Rough Sets

In this section, we will review some basic concepts such as knowledge base, Pawlak's rough set and Qian *et al.*'s multigranulation rough sets.

2.1 Pawlak's Rough Set

Let $U \neq \varnothing$ be a universe of discourse, A is a family of the equivalence relations on U, then the pair $K = (U, A)$ is referred as a knowledge base. If $P \subseteq A$ and $P \neq \varnothing$, then $\bigcap P$ (intersection of all equivalence relations in P) is also an equivalence relation, and will be denoted by $IND(P)$. It is referred as an indiscernibility relation over P in Pawlak's rough set theory.

$\forall R \in A$, we use U/R to represent the family of the equivalence classes, which are generated from the equivalence relation R. Therefore, $\forall x \in U$, $[x]_R$ is used to denote an equivalence class in terms of R, which contains x.

Suppose that $P \subseteq A$, then $IND(P)$ is also an equivalence relation, $U/IND(P)$ is then the family of the equivalence classes, which are generated from the equivalence relation $IND(P)$. Each element in $U/IND(P)$ is referred as a P-basic knowledge, $[x]_P = \{y \in U : (x, y) \in IND(P)\}$ is the equivalence class of $IND(P)$, which contains x. $\forall X \subseteq U$, if X is the union of some P-basic knowledge, then X is P-definable; otherwise X is P-undefinable. To describe the P-undefinable set more clearly, Pawlak proposed his rough set model as Definition 1 shows.

Definition 1[7]. *Let $K = (U, A)$ be a knowledge base, $P \subseteq A$, then $\forall X \subseteq U$, the lower approximation and upper approximation of X are denoted by $\underline{P}(X)$ and $\overline{P}(X)$, respectively,*

$$\underline{P}(X) = \{x \in U : [x]_P \subseteq X\}, \tag{1}$$
$$\overline{P}(X) = \{x \in U : [x]_P \cap X \neq \varnothing\}. \tag{2}$$

$[\underline{P}(X), \overline{P}(X)]$ is referred as the rough set of X.

2.2 Multigranulation Rough Sets

The multigranulation rough set approach is different from Pawlak's rough set approach because the former is constructed on the basis of a family of the binary relations instead of a single one. In Qian *et al.*'s multigranulation rough set theory, two different models have been defined. The first one is the optimistic multigranulation rough set, the second one is the pessimistic multigranulation rough set[16].

2.2.1 *Optimistic Multigranulation Rough Set*

Each equivalence relation can induce a partition on the universe of discourse, and such partition can be considered as a granulation space. Therefore, a family of equivalence relations can induce a family of granulation spaces. In optimistic multigranulation lower approximation, the word "optimistic" is used to express the idea that in multi–independent granulation spaces, we need only at least one of the granulation spaces to satisfy with the inclusion condition between the equivalence class and target. The upper approximation of optimistic multigranulation rough set is defined by the complement of the optimistic multigranulation lower approximation.

Definition 2[15]. *Let $K = (U, A)$ be a knowledge base, i.e., $\mathscr{R} = \{R_1, R_2, \ldots, R_m\}$ in which $R_1, R_2, \ldots, R_m \in A$, then $\forall X \subseteq U$, the optimistic multigranulation lower and upper approximations of X are denoted by $\underline{\mathscr{R}}^{\mathrm{OPT}}(X)$ and $\overline{\mathscr{R}}^{\mathrm{OPT}}(X)$, respectively,*

$$\underline{\mathscr{R}}^{\mathrm{OPT}}(X) = \{x \in U : \exists R_i \in \mathscr{R}, [x]_{R_i} \subseteq X\}, \tag{3}$$

$$\overline{\mathscr{R}}^{\mathrm{OPT}}(X) = \sim \underline{\mathscr{R}}^{\mathrm{OPT}}(\sim X), \tag{4}$$

where $\sim X$ is the complement of set X.

$[\underline{\mathscr{R}}^{\mathrm{OPT}}(X), \overline{\mathscr{R}}^{\mathrm{OPT}}(X)]$ is referred as the optimistic multigranulation rough set of X. By optimistic multigranulation lower and upper approximations, the optimistic multigranulation boundary region of X is

$$BN_{\mathscr{R}}^{\mathrm{OPT}}(X) = \overline{\mathscr{R}}^{\mathrm{OPT}}(X) - \underline{\mathscr{R}}^{\mathrm{OPT}}(X). \tag{5}$$

Theorem 1. *Let $K = (U, A)$ be a knowledge base, $\mathscr{R} = \{R_1, R_2, \ldots, R_m\}$ in which $R_1, R_2, \ldots, R_m \in A$, then $\forall X \subseteq U$, we have*

$$\overline{\mathscr{R}}^{\mathrm{OPT}}(X) = \{x \in U : \forall R_i \in \mathscr{R}, [x]_{R_i} \cap X \neq \varnothing\}. \tag{6}$$

Proof. It can be derived directly from Definition 2. □

By Theorem 1, we can see that though the optimistic multigranulation upper approximation is defined by the complement of the optimistic multigranulation lower approximation, it can also be considered as a set, in which objects have non–empty intersection with the target in terms of each granulation space.

2.2.2 *Pessimistic Multigranulation Rough Set*

In the pessimistic multigranulation rough set, the target is still approximated through a family of equivalence relations. However, the pessimistic case is different from the optimistic case. In pessimistic multigranulation lower approximation, the word "pessimistic" is used to express the idea that we need all of the granulation spaces to satisfy with the inclusion condition between the equivalence class and target. The upper approximation of pessimistic multigranulation rough set is still defined by the complement of the pessimistic multigranulation lower approximation.

Definition 3[16]**.** *Let $K = (U, A)$ be a knowledge base, $\mathscr{R} = \{R_1, R_2, \ldots, R_m\}$ in which $R_1, R_2, \ldots, R_m \in A$, then $\forall X \subseteq U$, the pessimistic multigranulation lower and upper approximations of X are denoted by $\underline{\mathscr{R}}^{\rm PES}(X)$ and $\overline{\mathscr{R}}^{\rm PES}(X)$, respectively,*

$$\underline{\mathscr{R}}^{\rm PES}(X) = \{x \in U : \forall R_i \in \mathscr{R}, [x]_{R_i} \subseteq X\}, \quad (7)$$

$$\overline{\mathscr{R}}^{\rm PES}(X) = \sim \underline{\mathscr{R}}^{\rm PES}(\sim X). \quad (8)$$

$[\underline{\mathscr{R}}^{\rm PES}(X), \overline{\mathscr{R}}^{\rm PES}(X)]$ is referred as the pessimistic multigranulation rough set of X. By pessimistic multigranulation lower and upper approximations, the pessimistic multigranulation boundary region of X is

$$BN^{\rm PES}_{\mathscr{R}}(X) = \overline{\mathscr{R}}^{\rm PES}(X) - \underline{\mathscr{R}}^{\rm PES}(X). \quad (9)$$

Theorem 2. *Let $K = (U, A)$ be a knowledge base, $\mathscr{R} = \{R_1, R_2, \ldots, R_m\}$ in which $R_1, R_2, \ldots, R_m \in A$, then $\forall X \subseteq U$, we have*

$$\overline{\mathscr{R}}^{\rm PES}(X) = \{x \in U : \exists R_i \in \mathscr{R}, [x]_{R_i} \cap X \neq \varnothing\}. \quad (10)$$

Proof. It can be derived directly from Definition 3. □

Different from upper approximation of optimistic multigranulation rough set, upper approximation of pessimistic multigranulation rough set is a set, in which objects have non–empty intersection with the target in terms of at least one of the granulation spaces.

For more details about multigranulation rough sets, we refer the readers to [13-16].

3 Hierarchical Structures on Partition-Based Multigranulation Spaces

Generally speaking, we can use a preference relation to represent the hierarchical structure on granulation spaces. However, it should be noticed that most of the proposed preference relations can only be used to judge whether a single granulation space is finer or coarser than another single granulation space, i.e., these preference relations are used to compare different single granulation spaces. In the above section, we have mentioned how to construct rough sets in multigranulation space and then, it is an interesting issue to discuss the hierarchical structures on multigranulation spaces.

Given a knowledge base, in which $R \in A$, since R is an equivalence relation, then we can induce a partition-based granulation space such as U/R. Moreover, suppose that $\mathscr{R} \subseteq A$, then the integration of all the granulation spaces forms a multigranulation space. Formally, a partition-based multigranulation space (PBMS) is denoted by $K(\mathscr{R})$ such that

$$K(\mathscr{R}) = \{U/R : R \in \mathscr{R}\}. \quad (11)$$

In (11), the multigranulation space is a family of the partitions on the universe of discourse. In the following, we will propose three different hierarchical structures to investigate the finer or coarser relations between two partition-based multigranulation spaces.

3.1 Definitions of Three Hierarchical Structures

Definition 4 (First Hierarchical Structure of PBMS). *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $\forall [x]_{R_j} (R_j \in \mathscr{R}_2)$, there must be $[x]_{R_k} (R_k \in \mathscr{R}_1)$ such that $[x]_{R_k} \subseteq [x]_{R_j}$, then we say that $K(\mathscr{R}_1)$ is finer than $K(\mathscr{R}_2)$ or $K(\mathscr{R}_2)$ is coarser than $K(\mathscr{R}_1)$, which is denoted by $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$ or $K(\mathscr{R}_2) \succeq_1 K(\mathscr{R}_1)$; if $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$ and $K(\mathscr{R}_1) \neq K(\mathscr{R}_2)$, then we say that $K(\mathscr{R}_1)$ is strictly finer than $K(\mathscr{R}_2)$, which is denoted by $K(\mathscr{R}_1) \prec_1 K(\mathscr{R}_2)$.*

Definition 5 (Second Hierarchical Structure of PBMS). *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $\forall [x]_{R_k} (R_k \in \mathscr{R}_1)$, there must be $[x]_{R_j} (R_j \in \mathscr{R}_2)$ such that $[x]_{R_k} \subseteq [x]_{R_j}$, then we say that $K(\mathscr{R}_1)$ is finer than $K(\mathscr{R}_2)$ or $K(\mathscr{R}_2)$ is coarser than $K(\mathscr{R}_1)$, which is denoted by $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$ or $K(\mathscr{R}_2) \succeq_2 K(\mathscr{R}_1)$; if $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$ and $K(\mathscr{R}_1) \neq K(\mathscr{R}_2)$, then we say that $K(\mathscr{R}_1)$ is strictly finer than $K(\mathscr{R}_2)$, which is denoted by $K(\mathscr{R}_1) \prec_2 K(\mathscr{R}_2)$.*

Definition 4 and Definition 5 are two different hierarchical structures, which are proposed to compare two partition-based multigranulation spaces. The first hierarchical structure says that each equivalence class in the second multigranulation space should include at least one of the equivalence classes in the first multigranulation space; the second hierarchical structure says that each equivalence class in the first multigranulation space should be included into at least one of the equivalence classes in the second multigranulation space. The following example will show that there is not a necessary causality between these two hierarchical structures.

Example 1. Suppose that $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ is the universe, three families of the equivalence relations are given by $K(\mathscr{R}_1) = \{U/R_1, U/R_2\}$, $K(\mathscr{R}_2) = \{U/R_3, U/R_4\}$ and $K(\mathscr{R}_3) = \{U/R_5, U/R_6\}$ such that $U/R_1 = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$, $U/R_2 = \{\{x_1, x_4, x_5\}, \{x_2, x_3\}, \{x_6, x_7, x_8\}\}$, $U/R_3 = \{\{x_1, x_2, x_3, x_4, x_5, x_6\}, \{x_7, x_8\}\}$, $U/R_4 = \{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$, $U/R_5 = \{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}\}$, $U/R_6 = \{\{x_1, x_2, x_3, x_7, x_8\}, \{x_4, x_6\}, \{x_5\}\}$.

By Definition 4, we can see that $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$ holds obviously. However, $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$ does not hold since $[x_6]_{R_2} \not\subseteq [x_6]_{R_3}$ and $[x_6]_{R_2} \not\subseteq [x_6]_{R_4}$.

By Definition 5, we can see that $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_3)$ holds obviously. However, $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_3)$ does not hold since $[x_4]_{R_1} \not\subseteq [x_4]_{R_6}$ and $[x_4]_{R_2} \not\subseteq [x_4]_{R_6}$.

Example 1 tells us that there is not a necessary causality from the fist hierarchical structure to second hierarchical structure; conversely, there is also not a necessary causality from the second hierarchical structure to the first hierarchical structure; i.e.,

$$\preceq_1 \not\longrightarrow \preceq_2, \quad \preceq_2 \not\longrightarrow \preceq_1 .$$

Theorem 3. *Let $K = (U, A)$ be a knowledge base, $\preceq_1$ is reflexive and transitive.*

Proof.

1) Suppose that $\mathscr{R} \subseteq A$, then by Definition 4, $K(\mathscr{R}) \preceq_1 K(\mathscr{R})$ holds obviously, $\preceq_1$ is reflexive.

2) Suppose that $\mathscr{R}_1, \mathscr{R}_2, \mathscr{R}_3 \subseteq A$, $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$ and $K(\mathscr{R}_2) \preceq_1 K(\mathscr{R}_3)$. By $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$ we know that $\forall [x]_{R_j} (R_j \in \mathscr{R}_2)$, there must be $[x]_{R_k} (R_k \in \mathscr{R}_1)$ such that $[x]_{R_k} \subseteq [x]_{R_j}$; by $K(\mathscr{R}_2) \preceq_1 K(\mathscr{R}_3)$ we know that $\forall [x]_{R_l} (R_l \in \mathscr{R}_3)$, there must be $[x]_{R_j} (R_j \in \mathscr{R}_2)$ such that $[x]_{R_j} \subseteq [x]_{R_l}$. Therefore, we can conclude that $\forall [x]_{R_l} (R_l \in \mathscr{R}_3)$, there must be $[x]_{R_k} (R_k \in \mathscr{R}_1)$ such that $[x]_{R_k} \subseteq [x]_{R_l}$, i.e., $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_3)$, it follows that $\preceq_1$ is transitive. □

Theorem 4. *Let $K = (U, A)$ be a knowledge base, $\preceq_2$ is reflexive and transitive.*

Proof. The proof of Theorem 4 is similar to the proof of Theorem 3. □

Example 2. Suppose that $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ is the universe, three families of the equivalence relations are given by $K(\mathscr{R}_1) = \{U/R_1, U/R_2\}$, $K(\mathscr{R}_2) = \{U/R_3, U/R_4\}$ and $K(\mathscr{R}_3) = \{U/R_5, U/R_6\}$ such that $U/R_1 = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_7, x_8\}\}$, $U/R_2 = \{\{x_1\}, \{x_2, x_3\}, \{x_4, x_5\}, \{x_6\}, \{x_7, x_8\}\}$, $U/R_3 = \{\{x_1, x_2, x_3\}, \{x_4, x_5\}, \{x_6\}, \{x_7, x_8\}\}$, $U/R_4 = \{\{x_1\}, \{x_2, x_3\}, \{x_4, x_5\}, \{x_6\}, \{x_7, x_8\}\}$, $U/R_5 = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4, x_5\}, \{x_6\}, \{x_7\}, \{x_8\}\}$, $U/R_6 = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_7, x_8\}\}$.

By Definition 4 and Definition 5, we can see that $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$ and $K(\mathscr{R}_2) \preceq_1 K(\mathscr{R}_1)$. However, $K(\mathscr{R}_1) \neq K(\mathscr{R}_2)$, from which we can see that $\preceq_1$ is a binary relation without the condition of the antisymmetric.

By Definition 4 and Definition 5, we can see that $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_3)$ and $K(\mathscr{R}_3) \preceq_2 K(\mathscr{R}_1)$. However, $K(\mathscr{R}_1) \neq K(\mathscr{R}_3)$, from which we can see that $\preceq_2$ is also a binary relation without the condition of the antisymmetric.

Example 2 tells us that both $\preceq_1$ and $\preceq_2$ are not antisymmetric. By such example and Theorem 3, we can conclude that such two hierarchical structures are not partial order in general.

By investigation of the above two hierarchical structures, it is not difficult to present the third hierarchical structure on partition-based multigranulation spaces as Definition 6 shows.

Definition 6 (Third Hierarchical Structure of PBMS). *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $\forall [x]_{R_k} (R_k \in \mathscr{R}_1)$ and $\forall [x]_{R_j} (R_j \in \mathscr{R}_2)$, we have $[x]_{R_k} \subseteq [x]_{R_j}$, then we say that $K(\mathscr{R}_1)$ is finer than $K(\mathscr{R}_2)$ or $K(\mathscr{R}_2)$ is coarser than $K(\mathscr{R}_1)$, which is denoted by $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$ or $K(\mathscr{R}_2) \succeq_3 K(\mathscr{R}_1)$; if $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$ and $K(\mathscr{R}_1) \neq K(\mathscr{R}_2)$, then we say that $K(\mathscr{R}_1)$ is strictly finer than $K(\mathscr{R}_2)$, which is denoted by $K(\mathscr{R}_1) \prec_3 K(\mathscr{R}_2)$.*

The third hierarchical structure says that each equivalence class in the first multigranulation space should be included into each equivalence class in the second multigranulation space in terms of each object in the universe. Obviously, the third hierarchical structure is stricter than both the first and the second hierarchical structures.

Theorem 5. *Let $K = (U, A)$ be a knowledge base, $\preceq_3$ is antisymmetric and transitive.*

Proof.

1) Suppose that $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$ and $K(\mathscr{R}_1) \neq K(\mathscr{R}_2)$, then there must be $[x]_{R_k} (R_k \in \mathscr{R}_1)$ and $[x]_{R_j} (R_j \in \mathscr{R}_2)$ such that $[x]_{R_k} \neq [x]_{R_j}$. It follows that $[x]_{R_k} \not\subseteq [x]_{R_j}$ or $[x]_{R_k} \not\supseteq [x]_{R_j}$. In other words, $\exists [x]_{R_k} (R_k \in \mathscr{R}_1)$ and $\exists [x]_{R_j} (R_j \in \mathscr{R}_2)$ such that $[x]_{R_k} \not\subseteq [x]_{R_j}$ or $[x]_{R_k} \not\supseteq [x]_{R_j}$, from which we can conclude that $K(\mathscr{R}_1) \not\preceq_3 K(\mathscr{R}_2)$ or $K(\mathscr{R}_2) \not\preceq_3 K(\mathscr{R}_1)$ and then $\preceq_3$ is antisymmetric.

2) Suppose that $\mathscr{R}_1, \mathscr{R}_2, \mathscr{R}_3 \subseteq A$, $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$ and $K(\mathscr{R}_2) \preceq_3 K(\mathscr{R}_3)$. By $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$ we know that $\forall [x]_{R_k} (R_k \in \mathscr{R}_1)$ and $\forall [x]_{R_j} (R_j \in \mathscr{R}_2)$, $[x]_{R_k} \subseteq [x]_{R_j}$ holds; by $K(\mathscr{R}_2) \preceq_3 K(\mathscr{R}_3)$ we know that $\forall [x]_{R_j} (R_j \in \mathscr{R}_2)$ and $\forall [x]_{R_l} (R_l \in \mathscr{R}_3)$, $[x]_{R_j} \subseteq [x]_{R_l}$ holds. Therefore, we can conclude that $\forall [x]_{R_k} (R_k \in \mathscr{R}_1)$ and $\forall [x]_{R_l} (R_l \in \mathscr{R}_3)$, $[x]_{R_k} \subseteq [x]_{R_l}$ holds, i.e., $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_3)$, $\preceq_3$ is transitive. □

It should be noticed that the third hierarchical structure is not necessarily reflexive if the multigranulation space contains two or more partitions. Nevertheless, if the multigranulation space degenerates into single granulation space, then the third hierarchical structure is reflexive.

Theorem 6. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$, then $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$ and $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$.*

Proof. By the above definitions of the three hierarchical structures, it is trivial to prove this theorem. □

The result in Theorem 6 is

$$\preceq_3 \longrightarrow \preceq_1 \quad \text{and} \quad \preceq_2 .$$

The following example will show that the inverse of Theorem 6 does not hold.

Example 3. Suppose that $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ is the universe, two families of the equivalence relations are given by $K(\mathscr{R}_1) = \{U/R_1, U/R_2\}$ and $K(\mathscr{R}_2) = \{U/R_3, U/R_4\}$ such that $U/R_1 = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_7, x_8\}\}$; $U/R_2 = \{\{x_1\}, \{x_2, x_3\}, \{x_4, x_5\}, \{x_6\}, \{x_7, x_8\}\}$; $U/R_3 = \{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}\}$; $U/R_4 = \{\{x_1\}, \{x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_7, x_8\}\}$.

By Definition 4 and Definition 5, $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$ and $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$ hold obviously. However, $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$ does not hold since $[x_1]_{R_1} \nsubseteq [x_1]_{R_4}$.

Example 3 tells us that there is not a necessary causality from the integration of the first and second hierarchical structures to the third hierarchical structure, i.e.,

$$\preceq_1 \quad \text{and} \quad \preceq_2 \not\longrightarrow \preceq_3 .$$

From what has been discussed in this subsection, the relationships among the proposed three hierarchical structures are

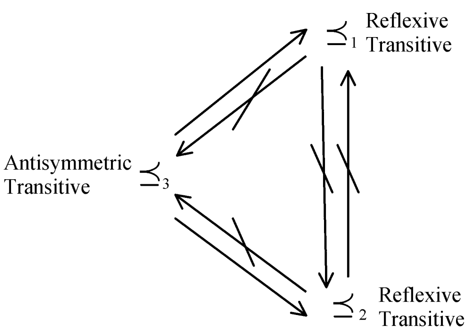

Fig.1. Relationships among the three hierarchical structures on partition-based multigranulation spaces.

3.2 Relationships between Hierarchical Structures and Multigranulation Rough Sets

In Section 2, optimistic and pessimistic multigranulation rough sets have been briefly introduced. In Subsection 3.1, three different hierarchical structures have been proposed. Then in the following, we will investigate the relationships between these hierarchical structures and the varieties of multigranulation rough sets.

Theorem 7. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$, then $\underline{\mathscr{R}_1}^{\mathrm{OPT}}(X) \supseteq \underline{\mathscr{R}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$.*

Proof. $\forall x \in \underline{\mathscr{R}_2}^{\mathrm{OPT}}(X)$, then by Definition 2, there must be $R_j \in \mathscr{R}_2$ such that $[x]_{R_j} \subseteq X$. By condition we know that $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$, then by Definition 4, there must be $[x]_{R_k} (R_k \in \mathscr{R}_1)$ such that $[x]_{R_k} \subseteq [x]_{R_j}$, and it follows that $[x]_{R_k} \subseteq X$, i.e., $x \in \underline{\mathscr{R}_1}^{\mathrm{OPT}}(X)$. □

Theorem 7 tells us that using the first hierarchical structure, if the partition-based multigranulation space is finer, then greater optimistic multigranulation lower approximation will be obtained.

Theorem 8. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$, then $\overline{\mathscr{R}_1}^{\mathrm{OPT}}(X) \subseteq \overline{\mathscr{R}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$.*

Proof. The proof of Theorem 8 is similar to the proof of Theorem 7. □

Theorem 8 tells us that using the first hierarchical structure, if the partition-based multigranulation space is finer, then the smaller optimistic multigranulation upper approximation will be obtained.

Theorem 9. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $\underline{\mathscr{R}_1}^{\mathrm{OPT}}(X) \supseteq \underline{\mathscr{R}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$, then $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$.*

Proof. Suppose that $K(\mathscr{R}_1) \npreceq_1 K(\mathscr{R}_2)$, then by Definition 4, $\exists [x]_{R_j} (R_j \in \mathscr{R}_2)$ such that $[x]_{R_k} \nsubseteq [x]_{R_j}$ for each $[x]_{R_k} (R_k \in \mathscr{R}_1)$. $x \in \underline{\mathscr{R}_2}^{\mathrm{OPT}}([x]_{R_j})$ holds obviously since $[x]_{R_j} \subseteq [x]_{R_j}$. However, $[x]_{R_k} \nsubseteq [x]_{R_j}$ holds for each $[x]_{R_k} (R_k \in \mathscr{R}_1)$, it follows that $x \notin \underline{\mathscr{R}_1}^{\mathrm{OPT}}([x]_{R_{2j}})$. From discussions above, we know that $\exists Y \subseteq U$ such that $\underline{\mathscr{R}_1}^{\mathrm{OPT}}(Y) \nsupseteq \underline{\mathscr{R}_2}^{\mathrm{OPT}}(Y)$. □

Theorem 9 is the inverse of Theorem 7, i.e., if the optimistic multigranulation lower approximation is greater, then the multigranulation space is finer in terms of the first hierarchical structure.

Theorem 10. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $\overline{\mathscr{R}_1}^{\mathrm{OPT}}(X) \subseteq \overline{\mathscr{R}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$, then $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$.*

Proof. The proof of Theorem 10 is similar to the proof of Theorem 9. □

Theorem 10 is the inverse of Theorem 8, i.e., if the optimistic multigranulation upper approximation is smaller, then the multigranulation space is finer in terms of the first hierarchical structure.

Example 4. Suppose that the director of the school must give a global evaluation to some students. This evaluation should be based on the level in mathematics, physics and literature. The director gave the examples of evaluation as shown in Table 1. The example contains eight students described by means of four attributes: $a_1 \Rightarrow$ level in mathematics (condition attribute), $a_2 \Rightarrow$ level in physics (condition attribute), $a_3 \Rightarrow$ level of literature (condition attribute), $d \Rightarrow$ global evaluation (decision attribute).

Table 1. Example of Students' Evaluations

U	a_1	a_2	a_3	d
x_1	2	3	2	Bad
x_2	5	1	3	Medium
x_3	5	2	4	Bad
x_4	3	5	3	Good
x_5	1	3	4	Bad
x_6	2	5	3	Medium
x_7	3	1	2	Bad
x_8	2	1	2	Medium

Suppose that $K(\mathscr{R}_1) = \{U/IND(\{a_1, a_2, a_3\})\}$, $K(\mathscr{R}_2) = \{U/IND(\{a_1\}), U/IND(\{a_2\}), U/IND(\{a_3\})\}$, then we can see that $K(\mathscr{R}_1) \preceq_1 K(\mathscr{R}_2)$ holds obviously. Since the decision attribute d partitions the universe into subsets such that $U/IND(\{d\}) = \{\text{Bad}, \text{Medium}, \text{Good}\} = \{\{x_1, x_3, x_5, x_7\}, \{x_2, x_6, x_8\}, \{x_4\}\}$, then it is not difficult to compute optimistic multigranulation rough sets in terms of two different multigranulation spaces as following:

$$\underline{\mathscr{R}_1}^{\text{OPT}}(\text{Bad}) = \{x_1, x_3, x_5, x_7\},$$
$$\underline{\mathscr{R}_1}^{\text{OPT}}(\text{Medium}) = \{x_2, x_6, x_8\},$$
$$\underline{\mathscr{R}_1}^{\text{OPT}}(\text{Good}) = \{x_4\};$$
$$\underline{\mathscr{R}_2}^{\text{OPT}}(\text{Bad}) = \{x_1, x_3, x_5\},$$
$$\underline{\mathscr{R}_2}^{\text{OPT}}(\text{Medium}) = \varnothing,$$
$$\underline{\mathscr{R}_2}^{\text{OPT}}(\text{Good}) = \varnothing;$$
$$\overline{\mathscr{R}_1}^{\text{OPT}}(\text{Bad}) = \{x_1, x_3, x_5, x_7\},$$
$$\overline{\mathscr{R}_1}^{\text{OPT}}(\text{Medium}) = \{x_2, x_6, x_8\},$$
$$\overline{\mathscr{R}_1}^{\text{OPT}}(\text{Good}) = \{x_4\};$$
$$\overline{\mathscr{R}_2}^{\text{OPT}}(\text{Bad}) = \{x_1, x_3, x_5, x_7, x_8\},$$
$$\overline{\mathscr{R}_2}^{\text{OPT}}(\text{Medium}) = \{x_2, x_6, x_8\},$$
$$\overline{\mathscr{R}_2}^{\text{OPT}}(\text{Good}) = \{x_4\}.$$

By the above results, we can see that $\underline{\mathscr{R}_1}^{\text{OPT}}(X) \supseteq \underline{\mathscr{R}_2}^{\text{OPT}}(X)$ and $\overline{\mathscr{R}_1}^{\text{OPT}}(X) \subseteq \overline{\mathscr{R}_2}^{\text{OPT}}(X)$ in which X is the decision class in $U/IND(\{d\})$. In other words, this example shows the relationships between the first hierarchical structure and optimistic multigranulation rough set in detail.

Theorem 11. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$, then $\underline{\mathscr{R}_1}^{\text{PES}}(X) \supseteq \underline{\mathscr{R}_2}^{\text{PES}}(X)$ for each $X \subseteq U$.*

Proof. $\forall x \in \underline{\mathscr{R}_2}^{\text{PES}}(X)$, by Definition 3, we know that $[x]_{R_j} \subseteq X$ holds for each $R_j \in \mathscr{R}_2$. By condition, $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$, then by Definition 5, $\forall [x]_{R_k} (R_k \in \mathscr{R}_1)$, there must be $[x]_{R_j} (R_j \in \mathscr{R}_2)$ such that $[x]_{R_k} \subseteq [x]_{R_j}$. Since $[x]_{R_j} \subseteq X$ for each $R_j \in \mathscr{R}_2$, then we can conclude that $[x]_{R_k} \subseteq X$ for each $R_k \in \mathscr{R}_1$, i.e., $x \in \underline{\mathscr{R}_1}^{\text{PES}}(X)$. □

Theorem 11 tells us that using the second hierarchical structure, if the partition-based multigranulation space is finer, then the greater pessimistic multigranulation lower approximation will be obtained.

Theorem 12. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$, then $\overline{\mathscr{R}_1}^{\text{PES}}(X) \subseteq \overline{\mathscr{R}_2}^{\text{PES}}(X)$ for each $X \subseteq U$.*

Proof. The proof of Theorem 12 is similar to the proof of Theorem 11. □

Theorem 12 tells us that using the second hierarchical structure, if the partition-based multigranulation space is finer, then the smaller pessimistic multigranulation upper approximation will be obtained.

Theorem 13. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $\underline{\mathscr{R}_1}^{\text{PES}}(X) \supseteq \underline{\mathscr{R}_2}^{\text{PES}}(X)$ for each $X \subseteq U$, then $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$.*

Proof. Suppose that $K(\mathscr{R}_1) \npreceq_2 K(\mathscr{R}_2)$, then by Definition 5, $\exists [x]_{R_k} (R_k \in \mathscr{R}_1)$ such that $[x]_{R_k} \nsubseteq [x]_{R_j}$ for each $[x]_{R_j} (R_j \in \mathscr{R}_2)$. Let $Y = \bigcup\{[x]_{R_j} : \forall R_j \in \mathscr{R}_2\}$, by Definition 3 we know that $x \in \underline{\mathscr{R}_2}^{\text{PES}}(Y)$ since $[x]_{R_j} \subseteq Y$ for each $R_j \in \mathscr{R}_2$. However, since $[x]_{R_k} \nsubseteq [x]_{R_j}$ for each $[x]_{R_j} (R_j \in \mathscr{R}_2)$, then $[x]_{R_k} \nsubseteq Y$, and it follows that $x \notin \underline{\mathscr{R}_1}^{\text{PES}}(Y)$. From discussion above, we know that $\exists Y \subseteq U$ such that $\underline{\mathscr{R}_1}^{\text{PES}}(Y) \nsupseteq \underline{\mathscr{R}_2}^{\text{PES}}(Y)$. □

Theorem 13 is the inverse of Theorem 11, i.e., if the pessimistic multigranulation lower approximation is greater, then the multigranulation space is finer in terms of the second hierarchical structure.

Theorem 14. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $\overline{\mathscr{R}_1}^{\text{PES}}(X) \subseteq \overline{\mathscr{R}_2}^{\text{PES}}(X)$ for each $X \subseteq U$, then $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$.*

Proof. The proof of Theorem 14 is similar to the proof of Theorem 13. □

Theorem 14 is the inverse of Theorem 12, i.e., if the pessimistic multigranulation upper approximation is smaller, then the multigranulation space is finer in terms of the second hierarchical structure.

Example 5. Take Table 1 for instance, by Definition 5, we know that $K(\mathscr{R}_1) \preceq_2 K(\mathscr{R}_2)$. Therefore, it is not difficult to obtain the following pessimistic multigranulation rough sets:

$$\underline{\mathscr{R}_1}^{\text{PES}}(\text{Bad}) = \{x_1, x_3, x_5, x_7\},$$
$$\underline{\mathscr{R}_1}^{\text{PES}}(\text{Medium}) = \{x_2, x_6, x_8\},$$
$$\underline{\mathscr{R}_1}^{\text{PES}}(\text{Good}) = \{x_4\};$$
$$\underline{\mathscr{R}_2}^{\text{PES}}(\text{Bad}) = \{x_5\},$$
$$\underline{\mathscr{R}_2}^{\text{PES}}(\text{Medium}) = \varnothing,$$
$$\underline{\mathscr{R}_2}^{\text{PES}}(\text{Good}) = \varnothing;$$
$$\overline{\mathscr{R}_1}^{\text{PES}}(\text{Bad}) = \{x_1, x_3, x_5, x_7\},$$

$\overline{\mathscr{R}_1}^{\mathrm{PES}}(\text{Medium}) = \{x_2, x_6, x_8\}$,
$\overline{\mathscr{R}_1}^{\mathrm{PES}}(\text{Good}) = \{x_4\}$;
$\overline{\mathscr{R}_2}^{\mathrm{PES}}(\text{Bad}) = U$,
$\overline{\mathscr{R}_2}^{\mathrm{PES}}(\text{Medium}) = \{x_1, x_2, x_3, x_4, x_6, x_7, x_8\}$,
$\overline{\mathscr{R}_2}^{\mathrm{PES}}(\text{Good}) = \{x_2, x_4, x_6, x_7\}$.

By the above results, we can see that $\underline{\mathscr{R}_1}^{\mathrm{PES}}(X) \supseteq \underline{\mathscr{R}_2}^{\mathrm{PES}}(X)$ and $\overline{\mathscr{R}_1}^{\mathrm{PES}}(X) \subseteq \overline{\mathscr{R}_2}^{\mathrm{PES}}(X)$ in which X is the decision class in $U/IND(\{d\})$. In other words, this example shows the relationships between the second hierarchical structure and pessimistic multigranulation rough set in detail.

Since the third hierarchical structure can generate the first and the second hierarchical structures (see Fig.1), then by the above theorems, it is not difficult to obtain the following corollaries.

Corollary 1. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$, then $\underline{\mathscr{R}_1}^{\mathrm{OPT}}(X) \supseteq \underline{\mathscr{R}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$.*

Corollary 2. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$, then $\overline{\mathscr{R}_1}^{\mathrm{OPT}}(X) \subseteq \overline{\mathscr{R}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$.*

Corollary 3. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$, then $\underline{\mathscr{R}_1}^{\mathrm{PES}}(X) \supseteq \underline{\mathscr{R}_2}^{\mathrm{PES}}(X)$ for each $X \subseteq U$.*

Corollary 4. *Let $K = (U, A)$ be a knowledge base in which $\mathscr{R}_1, \mathscr{R}_2 \subseteq A$, if $K(\mathscr{R}_1) \preceq_3 K(\mathscr{R}_2)$, then $\overline{\mathscr{R}_1}^{\mathrm{PES}}(X) \subseteq \overline{\mathscr{R}_2}^{\mathrm{PES}}(X)$ for each $X \subseteq U$.*

The above four corollaries tell us that if the partition-based multigranulation space is finer, then the optimistic and pessimistic multigranulation lower approximations are greater and the optimistic and pessimistic multigranulation upper approximations are smaller.

4 Multicovering Rough Sets

In Pawlak's knowledge base, an indiscernibility relation can induce a partition, which is the basic knowledge for approximating the target concept. However, in many practical applications, the granules are formed in the covering, in which two granules may not be disjoint to each other. In many rough set literatures, several types of covering-based rough sets have been proposed and deeply investigated. However, it should be noticed that those covering-based rough sets were constructed on the basis of one and only one covering. By employing the basic thinking of multigranulation rough sets, it is natural to study the multicovering rough sets approaches. To simplify our discussion, we will use a single covering-based rough set, which has been studied in [11, 30], to construct our multicovering rough sets.

4.1 Single Covering-Based Rough Set

Definition 7. *Let U be the universe of discourse, $B = \{C_1, C_2, \ldots, C_n\}$ is a family of the subsets of U, if $\bigcup B = U$, then B is referred as a covering on U.*

Definition 8. *Let U be the universe of discourse, $B = \{C_1, C_2, \ldots, C_n\}$ is a covering on U, $\forall X \subseteq U$, the lower and upper approximations of X can be denoted by $\underline{B}(X)$ and $\overline{B}(X)$, respectively,*

$$\underline{B}(X) = \{x \in U : N_B(x) \subseteq X\}, \tag{12}$$
$$\overline{B}(X) = \{x \in U : N_B(x) \cap X \neq \varnothing\}, \tag{13}$$

where $N_B(x) = \bigcap\{C : C \in B \wedge x \in C\}$ is the neighborhood of x.

4.2 Multicovering-Based Rough Sets

In the following, we may consider a family of the coverings on the universe and then propose the multicovering rough sets.

Definition 9. *Let U be the universe of discourse, $\mathscr{C} = \{B_1, B_2, \ldots, B_m\}$ is a family of the coverings on the universe of discourse, $\forall X \subseteq U$, the optimistic multicovering lower and upper approximations of X are denoted by $\underline{\mathscr{C}}^{\mathrm{OPT}}(X)$ and $\overline{\mathscr{C}}^{\mathrm{OPT}}(X)$, respectively,*

$$\underline{\mathscr{C}}^{\mathrm{OPT}}(X) = \{x \in U : \exists B_i \in \mathscr{C}, N_{B_i}(x) \subseteq X\}, \tag{14}$$

$$\overline{\mathscr{C}}^{\mathrm{OPT}}(X) = \sim \underline{\mathscr{C}}^{\mathrm{OPT}}(\sim X). \tag{15}$$

$[\underline{\mathscr{C}}^{\mathrm{OPT}}(X), \overline{\mathscr{C}}^{\mathrm{OPT}}(X)]$ is referred as the optimistic multicovering rough set of X. Similar to partition-based optimistic multigranulation rough set (see Definition 2), in multicovering case, each covering can induce a neighborhood for each object. If at least one of these neighborhoods for an object is included into the target concept, then such object belongs to the optimistic multicovering lower approximation. The optimistic multicovering upper approximation shows the complement property of the optimistic multicovering rough set.

Definition 10. *Let U be the universe of discourse, $\mathscr{C} = \{B_1, B_2, \ldots, B_m\}$ is a family of the coverings on the universe of discourse, $\forall X \subseteq U$, the pessimistic multicovering lower and upper approximations of X are denoted by $\underline{\mathscr{C}}^{\mathrm{PES}}(X)$ and $\overline{\mathscr{C}}^{\mathrm{PES}}(X)$, respectively,*

$$\underline{\mathscr{C}}^{\mathrm{PES}}(X) = \{x \in U : \forall B_i \in \mathscr{C}, N_{B_i}(x) \subseteq X\}, \tag{16}$$

$$\overline{\mathscr{C}}^{\mathrm{PES}}(X) = \sim \underline{\mathscr{C}}^{\mathrm{PES}}(\sim X). \tag{17}$$

$[\underline{\mathscr{C}}^{\mathrm{PES}}(X), \overline{\mathscr{C}}^{\mathrm{PES}}(X)]$ is referred as the pessmistic multicovering rough set of X. Similar to partition-based pessimistic multigranulation rough set (see Definition 3), in multicovering case, each covering can induce a neighborhood for each object. If all of these neighborhoods of an object is included into the target concept, then such object belongs to the pessimistic multicovering lower approximation. The pessimistic multicovering upper approximation shows the complement property of the pessimistic multicovering rough set.

Theorem 15. *Let U be the universe of discourse, $\mathscr{C} = \{B_1, B_2, \ldots, B_m\}$ is a family of the coverings on the universe of discourse, $\forall X \subseteq U$, we have*

$$\overline{\mathscr{C}}^{\mathrm{OPT}}(X) = \{x \in U : \forall B_i \in \mathscr{C}, N_{B_i}(x) \cap X \neq \varnothing\}. \tag{18}$$

Theorem 16. *Let U be the universe of discourse, $\mathscr{C} = \{B_1, B_2, \ldots, B_m\}$ is a family of the coverings on the universe of discourse, $\forall X \subseteq U$, we have*

$$\overline{\mathscr{C}}^{\mathrm{PES}}(X) = \{x \in U : \exists B_i \in \mathscr{C}, N_{B_i}(x) \cap X \neq \varnothing\}. \tag{19}$$

Theorem 15 tells us that though optimistic multicovering upper approximation is defined by the complement of optimistic multicovering lower approximation, it can also be considered as a set, in which objects have non-empty intersection with the target in terms of each covering. Theorem 16 tells us that though pessimistic multicovering upper approximation is defined by the complement of pessimistic multicovering lower approximation, it can also be considered as a set, in which objects have non-empty intersection with the target in terms of at least of one of the coverings.

Proposition 1. *Let U be the universe of discourse, $\mathscr{C} = \{B_1, B_2, \ldots, B_m\}$ is a family of the coverings on the universe of discourse, then we have following properties about optimistic multicovering lower and upper approximations*:

1) $\underline{\mathscr{C}}^{\mathrm{OPT}}(X) \subseteq X \subseteq \overline{\mathscr{C}}^{\mathrm{OPT}}(X)$,

2) $\underline{\mathscr{C}}^{\mathrm{OPT}}(\varnothing) = \overline{\mathscr{C}}^{\mathrm{OPT}}(\varnothing) = \varnothing$,
$\underline{\mathscr{C}}^{\mathrm{OPT}}(U) = \overline{\mathscr{C}}^{\mathrm{OPT}}(U) = U$,

3) $\underline{\mathscr{C}}^{\mathrm{OPT}}(X) = \bigcup_{i=1}^{m} \underline{B_i}(X)$,
$\overline{\mathscr{C}}^{\mathrm{OPT}}(X) = \bigcap_{i=1}^{m} \overline{B_i}(X)$,

4) $X \subseteq Y \Rightarrow \underline{\mathscr{C}}^{\mathrm{OPT}}(X) \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(Y)$,
$X \subseteq Y \Rightarrow \overline{\mathscr{C}}^{\mathrm{OPT}}(X) \subseteq \overline{\mathscr{C}}^{\mathrm{OPT}}(Y)$,

5) $\underline{\mathscr{C}}^{\mathrm{OPT}}\left(\underline{\mathscr{C}}^{\mathrm{OPT}}(X)\right) = \underline{\mathscr{C}}^{\mathrm{OPT}}(X)$,
$\overline{\mathscr{C}}^{\mathrm{OPT}}\left(\overline{\mathscr{C}}^{\mathrm{OPT}}(X)\right) = \overline{\mathscr{C}}^{\mathrm{OPT}}(X)$.

Proof.

1) $\forall x \in \underline{\mathscr{C}}^{\mathrm{OPT}}(X)$, by Definition 9, there must be $B_i \in \mathscr{C}$ such that $N_{B_i}(x) \subseteq X$. Since $N_{B_i}(x) = \bigcap\{C : C \in B_i \wedge x \in C\}$, then we have $x \in N_{B_i}(x)$, and it follows that $x \in X$, i.e., $\underline{\mathscr{C}}^{\mathrm{OPT}}(X) \subseteq X$.

$\forall x \notin \overline{\mathscr{C}}^{\mathrm{OPT}}(X)$, by (15), we know that $x \in \underline{\mathscr{C}}^{\mathrm{OPT}}(\sim X)$. In other words, there must be $B_i \in \mathscr{C}$ such that $N_{B_i}(x) \subseteq \sim X$. Since $x \in N_{B_i}(x)$, then $x \in \sim X$, $x \notin X$, from which we can conclude that $X \subseteq \overline{\mathscr{C}}^{\mathrm{OPT}}(X)$.

2) By the result of 1) we know that $\underline{\mathscr{C}}^{\mathrm{OPT}}(\varnothing) \subseteq \varnothing$. Thus, it must be proved that $\varnothing \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(\varnothing)$. For each $x \notin \underline{\mathscr{C}}^{\mathrm{OPT}}(\varnothing)$, by Theorem 15, $\forall B_i \in \mathscr{C}$, $N_{B_i}(x) \nsubseteq \varnothing$ holds, and it follows that $x \in U$, i.e., $x \notin \varnothing$, from which we can conclude that $\varnothing \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(\varnothing)$.

Similarly, it is not difficult to prove that $\overline{\mathscr{C}}^{\mathrm{OPT}}(\varnothing) = \varnothing$.

By the result of 1) we know that $\underline{\mathscr{C}}^{\mathrm{OPT}}(U) \subseteq U$. Thus, it must be proved that $U \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(U)$. For each $x \notin \underline{\mathscr{C}}^{\mathrm{OPT}}(U)$, by Theorem 15, $\forall B_i \in \mathscr{C}$, $N_{B_i}(x) \nsubseteq U$ holds, and it follows that $x \in \varnothing$, i.e., $x \notin U$, from which we can conclude that $U \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(U)$.

Similarly, it is not difficult to prove that $\overline{\mathscr{C}}^{\mathrm{OPT}}(U) = U$.

3) $\forall x \in U$, by Definition 9, we have

$$\begin{aligned} x \in \underline{\mathscr{C}}^{\mathrm{OPT}}(X) &\Leftrightarrow \exists B_i \in \mathscr{C} \text{ s.t. } N_{B_i}(x) \subseteq X \\ &\Leftrightarrow \exists B_i \in \mathscr{C} \text{ s.t. } x \in \underline{B_i}(X) \\ &\Leftrightarrow x \in \textstyle\bigcup_{i=1}^{m} \underline{B_i}(X). \end{aligned}$$

$\forall x \in U$, by Theorem 15, we have

$$\begin{aligned} x \in \overline{\mathscr{C}}^{\mathrm{OPT}}(X) &\Leftrightarrow \forall B_i \in \mathscr{C}, N_{B_i}(x) \cap X \neq \varnothing \\ &\Leftrightarrow \forall B_i \in \mathscr{C}, \quad x \in \overline{B_i}(X) \\ &\Leftrightarrow x \in \textstyle\bigcap_{i=1}^{m} \overline{B_i}(X). \end{aligned}$$

4) $\forall x \in \underline{\mathscr{C}}^{\mathrm{OPT}}(X)$, there must be $B_i \in \mathscr{C}$ such that $N_{B_i}(x) \subseteq X$. Since $X \subseteq Y$, then $N_{B_i}(x) \subseteq Y$ also holds, it follows that $x \in \underline{\mathscr{C}}^{\mathrm{OPT}}(Y)$, i.e., $\underline{\mathscr{C}}^{\mathrm{OPT}}(X) \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(Y)$.

Similarly, it is not difficult to prove $\overline{\mathscr{C}}^{\mathrm{OPT}}(X) \subseteq \overline{\mathscr{C}}^{\mathrm{OPT}}(Y)$.

5) By the result of 1), we have $\underline{\mathscr{C}}^{\mathrm{OPT}}\left(\underline{\mathscr{C}}^{\mathrm{OPT}}(X)\right) \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(X)$. Thus, it must be proved that $\underline{\mathscr{C}}^{\mathrm{OPT}}\left(\underline{\mathscr{C}}^{\mathrm{OPT}}(X)\right) \supseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(X)$.

$\forall x \in \underline{\mathscr{C}}^{\mathrm{OPT}}(X)$, there must be $B_i \in \mathscr{C}$ such that $N_{B_i}(x) \subseteq X$. By the result of 4), we have $\underline{\mathscr{C}}^{\mathrm{OPT}}(N_{B_i}(x)) \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(X)$. $\forall y \in N_{B_i}(x)$, we have $N_{B_i}(y) \subseteq N_{B_i}(x)$ and then $y \in \underline{\mathscr{C}}^{\mathrm{OPT}}(N_{B_i}(x))$, from which we can conclude that $N_{B_i}(x) \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(N_{B_i}(x)) \subseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(X)$, i.e., $x \in \underline{\mathscr{C}}^{\mathrm{OPT}}\left(\underline{\mathscr{C}}^{\mathrm{OPT}}(X)\right)$. From discussions above, $\underline{\mathscr{C}}^{\mathrm{OPT}}\left(\underline{\mathscr{C}}^{\mathrm{OPT}}(X)\right) \supseteq \underline{\mathscr{C}}^{\mathrm{OPT}}(X)$ holds.

$\forall X \subseteq U$, by the above results, we have $\underline{\mathscr{C}}^{\mathrm{OPT}}\left(\underline{\mathscr{C}}^{\mathrm{OPT}}(\sim X)\right) = \underline{\mathscr{C}}^{\mathrm{OPT}}(\sim X)$. By Definition

9, we have the following:

$$\underline{\mathscr{C}}^{\rm OPT}\Big(\underline{\mathscr{C}}^{\rm OPT}(\sim X)\Big)=\underline{\mathscr{C}}^{\rm OPT}(\sim X)$$
$$\Rightarrow \underline{\mathscr{C}}^{\rm OPT}\Big(\sim \overline{\mathscr{C}}^{\rm OPT}(X)\Big)=\sim \overline{\mathscr{C}}^{\rm OPT}(X)$$
$$\Rightarrow \sim \overline{\mathscr{C}}^{\rm OPT}\Big(\overline{\mathscr{C}}^{\rm OPT}(X)\Big)=\sim \overline{\mathscr{C}}^{\rm OPT}(X)$$
$$\Rightarrow \overline{\mathscr{C}}^{\rm OPT}\Big(\overline{\mathscr{C}}^{\rm OPT}(X)\Big)=\overline{\mathscr{C}}^{\rm OPT}(X).$$

That completes the proof of $\overline{\mathscr{C}}^{\rm OPT}\Big(\overline{\mathscr{C}}^{\rm OPT}(X)\Big) = \overline{\mathscr{C}}^{\rm OPT}(X)$. □

Proposition 2. *Let U be the universe of discourse, $\mathscr{C}=\{B_1,B_2,\ldots,B_m\}$ is a family of the coverings on the universe of discourse, then we have following properties about pessimistic multicovering lower and upper approximations*:

1) $\underline{\mathscr{C}}^{\rm PES}(X)\subseteq X\subseteq \overline{\mathscr{C}}^{\rm PES}(X)$,

2) $\underline{\mathscr{C}}^{\rm PES}(\varnothing)=\overline{\mathscr{C}}^{\rm PES}(\varnothing)=\varnothing$,
$\underline{\mathscr{C}}^{\rm PES}(U)=\overline{\mathscr{C}}^{\rm PES}(U)=U$,

3) $\underline{\mathscr{C}}^{\rm PES}(X)=\bigcap_{i=1}^{m}\underline{B_i}(X)$,
$\overline{\mathscr{C}}^{\rm PES}(X)=\bigcup_{i=1}^{m}\overline{B_i}(X)$,

4) $X\subseteq Y\Rightarrow \underline{\mathscr{C}}^{\rm PES}(X)\subseteq \underline{\mathscr{C}}^{\rm PES}(Y)$,
$X\subseteq Y\Rightarrow \overline{\mathscr{C}}^{\rm PES}(X)\subseteq \overline{e\mathscr{C}}^{\rm PES}(Y)$,

5) $\underline{\mathscr{C}}^{\rm PES}\Big(\underline{\mathscr{C}}^{\rm PES}(X)\Big)=\underline{\mathscr{C}}^{\rm PES}(X)$,
$\overline{\mathscr{C}}^{\rm PES}\Big(\overline{\mathscr{C}}^{\rm PES}(X)\Big)=\overline{\mathscr{C}}^{\rm PES}(X)$.

Proof. The proof of Proposition 2 is similar to the proof of Proposition 1. □

5 Hierarchical Structures on Covering-Based Multigranulation Spaces

In this section, we may generalize the hierarchical structures, which have been proposed in Section 3 into covering-based multigranulation spaces (CBMSs).

Let U be the universe of discourse, B is a covering on U. Moreover, suppose that $\mathscr{C}=\{B_1,B_2,\ldots,B_m\}$ is a family of the coverings on the universe of discourse, then the integration of these coverings forms a covering-based multigranulation space. Formally, a covering-based multigranulation space is denoted by $K(\mathscr{C})$ where

$$K(\mathscr{C})=\{B_i:B_i\in\mathscr{C}\}. \tag{20}$$

5.1 Definitions of Three Hierarchical Structures

Definition 11 (First Hierarchical Structure of CBMS). *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $\forall N_{B_j}(x)(B_j\in\mathscr{C}_2)$, there must be $N_{B_k}(x)(B_k\in\mathscr{C}_1)$ such that $N_{B_k}(x)\subseteq N_{B_j}(x)$, then we say that $K(\mathscr{C}_1)$ is finer than $K(\mathscr{C}_2)$ or $K(\mathscr{C}_2)$ is coarser than $K(\mathscr{C}_1)$, which is denoted by $K(\mathscr{C}_1)\preceq_1 K(\mathscr{C}_2)$ or $K(\mathscr{C}_2)\succeq_1 K(\mathscr{C}_1)$; if $K(\mathscr{C}_1)\preceq_1 K(\mathscr{C}_2)$ and $K(\mathscr{C}_1)\neq K(\mathscr{C}_2)$, then we say that $K(\mathscr{C}_1)$ is strictly finer than $K(\mathscr{C}_2)$, which is denoted by $K(\mathscr{C}_1)\prec_1 K(\mathscr{C}_2)$.*

Definition 12 (Second Hierarchical Structure of CBMS). *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $\forall N_{B_k}(x)(B_k\in\mathscr{C}_1)$, there must be $N_{B_j}(x)(B_j\in\mathscr{C}_2)$ such that $N_{B_k}(x)\subseteq N_{B_j}(x)$, then we say that $K(\mathscr{C}_1)$ is finer than $K(\mathscr{C}_2)$ or $K(\mathscr{C}_2)$ is coarser than $K(\mathscr{C}_1)$, which is denoted by $K(\mathscr{C}_1)\preceq_2 K(\mathscr{C}_2)$ or $K(\mathscr{C}_2)\succeq_2 K(\mathscr{C}_1)$; if $K(\mathscr{C}_1)\preceq_2 K(\mathscr{C}_2)$ and $K(\mathscr{C}_1)\neq K(\mathscr{C}_2)$, then we say that $K(\mathscr{C}_1)$ is strictly finer than $K(\mathscr{C}_2)$, which is denoted by $K(\mathscr{C}_1)\prec_2 K(\mathscr{C}_2)$.*

The explanations of above two hierarchical structures are similar to those on PBMSs.

Example 6. Suppose that $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8\}$ is the universe, three families of the coverings are given by $\mathscr{C}_1=\{B_1,B_2\}$, $\mathscr{C}_2=\{B_3,B_4\}$ and $\mathscr{C}_3=\{B_5,B_6\}$ such that $B_1=\{\{x_1,x_2,x_3,x_4\},\{x_3,x_4,x_6\},\{x_5,x_6,x_7\},\{x_7,x_8\}\}$, $B_2=\{\{x_1,x_2,x_4,x_5,x_6\},\{x_2,x_3,x_4,x_6\},\{x_5,x_6,x_7,x_8\}\}$, $B_3=\{\{x_1,x_2,x_3,x_4,x_5,x_6\},\{x_5,x_6,x_7,x_8\}\}$, $B_4=\{\{x_1,x_2,x_3,x_4\},\{x_7,x_8\},\{x_3,x_4,x_5,x_6\},\{x_5,x_6,x_7,x_8\}\}$, $B_5=\{\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8\},\{x_5,x_6,x_7,x_8\}\}$, $B_6=\{\{x_1,x_2,x_3\},\{x_4,x_7,x_8\},\{x_4,x_6\},\{x_5\}\}$.

By Definition 11, we can see that $K(\mathscr{C}_1)\preceq_1 K(\mathscr{C}_2)$ holds obviously. However, $K(\mathscr{C}_1)\preceq_2 K(\mathscr{C}_2)$ does not hold since $N_{B_1}(x_5)\nsubseteq N_{B_3}(x_5)$ and $N_{B_1}(x_5)\nsubseteq N_{B_4}(x_5)$.

By Definition 12, we can see that $K(\mathscr{C}_1)\preceq_2 K(\mathscr{C}_3)$ holds obviously. However, $K(\mathscr{C}_1)\preceq_1 K(\mathscr{C}_3)$ does not hold since $N_{B_1}(x_1)\nsubseteq N_{B_6}(x_1)$ and $N_{B_2}(x_1)\nsubseteq N_{B_6}(x_1)$.

Example 6 tells us that there is not a necessary causality from the fist hierarchical structure to the second hierarchical structure; conversely, there is also not a necessary causality from the second hierarchical structure to the first hierarchical structure. These results are same with the partition case.

Theorem 17. *Let U be the universe of discourse, $\preceq_1$ is reflexive and transitive.*

Proof. The proof of Theorem 17 is similar to the proof of Theorem 3. □

Theorem 18. *Let U be the universe of discourse, $\preceq_2$ is reflexive and transitive.*

Proof. The proof of Theorem 18 is similar to the proof of Theorem 3. □

Example 7. Suppose that $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8\}$ is the universe, three families of the coverings are given by $\mathscr{C}_1=\{B_1,B_2\}$, $\mathscr{C}_2=\{B_3,$

$B_4\}$ and $\mathscr{C}_3 = \{B_5, B_6\}$ such that $B_1 = \{\{x_1, x_2, x_3, x_4, x_5\}, \{x_4, x_5, x_6\}, \{x_4, x_5, x_6, x_7, x_8\}\}$, $B_2 = \{\{x_1\}, \{x_2, x_3\}, \{x_4, x_5\}, \{x_6\}, \{x_7, x_8\}\}$, $B_3 = \{\{x_1, x_2, x_3, x_4, x_5\}, \{x_4, x_5\}, \{x_4, x_5, x_6\}, \{x_7, x_8\}\}$, $B_4 = \{\{x_1\}, \{x_1, x_2, x_3\}, \{x_2, x_3, x_4, x_5\}, \{x_6\}, \{x_6, x_7, x_8\}, \{x_7, x_8\}\}$, $B_5 = \{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4, x_5\}, \{x_6\}, \{x_4, x_5, x_6, x_7, x_8\}\}$, $B_6 = \{\{x_1, x_2, x_3, x_4, x_5\}, \{x_4, x_5, x_6\}, \{x_7, x_8\}\}$.

By Definition 11, we can see that $K(\mathscr{C}_1) \preceq_1 K(\mathscr{C}_2)$ and $K(\mathscr{C}_2) \preceq_1 K(\mathscr{C}_1)$. However, $K(\mathscr{C}_1) \neq K(\mathscr{C}_2)$, from which we can see that $\preceq_1$ is a binary relation without the condition of the antisymmetric.

By Definition 12, we can see that $K(\mathscr{C}_1) \preceq_2 K(\mathscr{C}_3)$ and $K(\mathscr{C}_3) \preceq_2 K(\mathscr{C}_1)$. However, $K(\mathscr{C}_1) \neq K(\mathscr{C}_3)$, from which we can see that $\preceq_2$ is also a binary relation without the condition of the antisymmetric.

Example 7 shows that the first and the second hierarchical structures on CBMSs are also antisymmetric.

Similar to the partition case, we may define the third hierarchical structure on CBMS as Definition 13 shows.

Definition 13 (Third Hierarchical Structure of CBMS). *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $\forall N_{B_j}(x)(B_j \in \mathscr{C}_2)$ and $\forall N_{B_k}(x)(B_k \in \mathscr{C}_1)$, we have $N_{B_k}(x) \subseteq N_{B_j}(x)$, then we say that $K(\mathscr{C}_1)$ is finer than $K(\mathscr{C}_2)$ or $K(\mathscr{C}_2)$ is coarser than $K(\mathscr{C}_1)$, which is denoted by $K(\mathscr{C}_1) \preceq_3 K(\mathscr{C}_2)$ or $K(\mathscr{C}_2) \succeq_3 K(\mathscr{C}_1)$; if $K(\mathscr{C}_1) \preceq_3 K(\mathscr{C}_2)$ and $K(\mathscr{C}_1) \neq K(\mathscr{C}_2)$, then we say that $K(\mathscr{C}_1)$ is strictly finer than $K(\mathscr{C}_2)$, which is denoted by $K(\mathscr{C}_1) \prec_3 K(\mathscr{C}_2)$.*

Theorem 19. *Let U be the universe of discourse, $\preceq_3$ is antisymmetric and transitive.*

Proof. The proof of Theorem 19 is similar to the proof of Theorem 5. □

Theorem 20. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $K(\mathscr{C}_1) \preceq_3 K(\mathscr{C}_2)$, then $K(\mathscr{C}_1) \preceq_1 K(\mathscr{C}_2)$ and $K(\mathscr{C}_1) \preceq_2 K(\mathscr{C}_2)$.*

Proof. By the definitions of the above three hierarchical structures, it is a trivial to prove this theorem. □

The following example will show that the inverse of Theorem 20 does not hold.

Example 8. Suppose that $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ is the universe, two families of the coverings are given by $\mathscr{C}_1 = \{B_1, B_2\}$ and $\mathscr{C}_2 = \{B_3, B_4\}$ such that $B_1 = \{\{x_1, x_2, x_3\}, \{x_2, x_3, x_4, x_5, x_6\}, \{x_7, x_8\}\}$, $B_2 = \{\{x_1\}, \{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}, \{x_6\}, \{x_7, x_8\}\}$, $B_3 = \{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}\}$, $B_4 = \{\{x_1\}, \{x_2, x_3\}, \{x_2, x_3, x_4, x_5, x_6\}, \{x_2, x_3, x_7, x_8\}\}$.

By Definition 11 and Definition 12, $K(\mathscr{C}_1) \preceq_1 K(\mathscr{C}_2)$ and $K(\mathscr{C}_1) \preceq_2 K(\mathscr{C}_2)$ hold obviously. However, $K(\mathscr{C}_1) \preceq_3 K(\mathscr{C}_2)$ does not hold since $N_{B_2}(x_2) \nsubseteq N_{B_4}(x_2)$.

Example 8 tells us that in CBMSs, there is not a necessary causality from the integration of the first and second hierarchical structures to the third hierarchical structure.

From discussions above, we may also obtain the relationships among the three hierarchical structures on CBMSs (see Fig.2), which are same to the partition case.

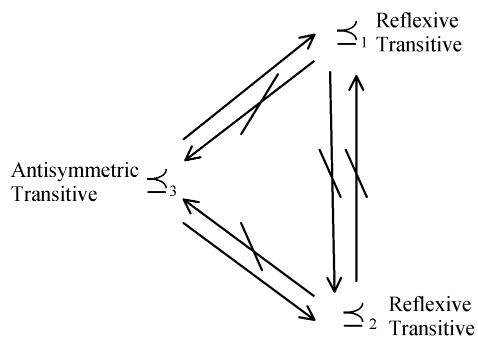

Fig.2. Relationships among three hierarchical structures on covering-based multigranulation spaces.

5.2 Relationships between Hierarchical Structures and Multicovering Rough Sets

Theorem 21. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $K(\mathscr{C}_1) \preceq_1 K(\mathscr{C}_2)$, then $\underline{\mathscr{C}_1}^{\mathrm{OPT}}(X) \supseteq \underline{\mathscr{C}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$.*

Proof. The proof of Theorem 21 is similar to the proof of Theorem 7. □

Theorem 21 tells us that using the first hierarchical structure, if the CBMS is finer, then the greater optimistic multicovering lower approximation will be obtained.

Theorem 22. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $K(\mathscr{C}_1) \preceq_1 K(\mathscr{C}_2)$, then $\overline{\mathscr{C}_1}^{\mathrm{OPT}}(X) \subseteq \overline{\mathscr{C}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$.*

Proof. The proof of Theorem 21 is similar to the proof of Theorem 7. □

Theorem 22 tells us that using the first hierarchical structure, if the CBMS is finer, then the smaller optimistic multicovering upper approximation will be obtained.

Theorem 23. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $\underline{\mathscr{C}_1}^{\mathrm{OPT}}(X) \supseteq \underline{\mathscr{C}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$, then $K(\mathscr{C}_1) \preceq_1 K(\mathscr{C}_2)$.*

Proof. The proof of Theorem 23 is similar to the proof of Theorem 9. □

Theorem 23 is the inverse of Theorem 21, i.e., if the optimistic multicovering lower approximation is greater, then the multigranulation space is finer in terms of the first hierarchical structure.

Theorem 24. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $\overline{\mathscr{C}_1}^{\rm OPT}(X) \subseteq \overline{\mathscr{C}_2}^{\rm OPT}(X)$ for each $X \subseteq U$, then $K(\mathscr{C}_1) \preceq_1 K(\mathscr{C}_2)$.*

Proof. The proof of Theorem 24 is similar to the proof of Theorem 9. □

Theorem 24 is the inverse of Theorem 22, i.e., if the optimistic multicovering upper approximation is smaller, then the multigranulation space is finer in terms of the first hierarchical structure.

Example 9. Let us consider an incomplete decision system for students' evaluations, which is shown in Table 2. Similar to Table 1, in Table 2, 15 students are evaluated by four attributes: $a_1 \Rightarrow$ level in mathematics (condition attribute), $a_2 \Rightarrow$ level in physics (condition attribute), $a_3 \Rightarrow$ level of literature (condition attribute), $d \Rightarrow$ global evaluation (decision attribute).

Table 2. Example of Incomplete Students' Evaluations

U	a_1	a_2	a_3	d
x_1	Medium	Medium	Bad	Bbad
x_2	Good	Medium	Bad	Medium
x_3	Bad	Good	*	Bad
x_4	Medium	Good	Bad	Medium
x_5	*	Good	Bad	Medium
x_6	Bad	Bad	Medium	Bad
x_7	Good	Bad	Medium	Bad
x_8	Medium	*	Medium	Medium
x_9	Good	Medium	Medium	Good
x_{10}	Medium	Good	Medium	Good
x_{11}	Good	Bad	Good	Bad
x_{12}	*	Medium	*	Medium
x_{13}	Good	Medium	Good	Good
x_{14}	Bad	Good	Good	Bad
x_{15}	Medium	Good	Good	Good

Since $*$ is used to denote unknown value in incomplete information system, then by the maximal consistent block technique, which was proposed in [31], we can obtain the following coverings:

1) set of maximal consistent blocks in terms of set of attributes $\{a_1, a_2, a_3\}$: $B_1 = \{\{x_1, x_{12}\}, \{x_2, x_{12}\}, \{x_9, x_{12}\}, \{x_{15}, x_{12}\}, \{x_{13}, x_{12}\}, \{x_8, x_{10}\}, \{x_8, x_{12}\}, \{x_3, x_{14}\}, \{x_3, x_5\}, \{x_4, x_5\}, \{x_6\}, \{x_7\}, \{x_{11}\}\}$;

2) set of maximal consistent blocks in terms of attribute a_1: $B_2 = \{\{x_1, x_4, x_5, x_8, x_{10}, x_{12}, x_{15}\}, \{x_2, x_5, x_7, x_9, x_{11}, x_{12}, x_{13}\}, \{x_3, x_5, x_6, x_{12}, x_{14}\}\}$;

3) set of maximal consistent blocks in terms of attribute a_2: $B_3 = \{\{x_1, x_2, x_8, x_9, x_{12}, x_{13}\}, \{x_3, x_4, x_5, x_8, x_{10}, x_{14}, x_{15}\}, \{x_6, x_7, x_8, x_{11}\}\}$;

4) set of maximal consistent blocks in terms of attribute a_3: $B_4 = \{\{x_1, x_2, x_3, x_4, x_5, x_{12}\}, \{x_3, x_6, x_7, x_8, x_9, x_{10}, x_{12}\}, \{x_3, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}\}\}$.

Suppose that $K(\mathscr{C}_1) = \{B_1\}$, $K(\mathscr{C}_2) = \{B_2, B_3, B_4\}$, then $K(\mathscr{C}_1) \preceq_1 K(\mathscr{C}_2)$ holds. Since d partitions the universe into subsets such that $U/IND(\{d\}) = \{\text{Bad}, \text{Medium}, \text{Good}\} = \{\{x_1, x_3, x_6, x_7, x_{11}, x_{14}\}, \{x_2, x_4, x_5, x_8, x_{12}\}, \{x_9, x_{10}, x_{13}, x_{15}\}\}$, then it is not difficult to compute the optimistic multicovering rough sets in terms of $K(\mathscr{C}_1)$ and $K(\mathscr{C}_2)$ as following:

$$\begin{aligned}
&\underline{\mathscr{C}_1}^{\rm OPT}(\text{Bad}) = \{x_3, x_6, x_7, x_{11}, x_{14}\},\\
&\underline{\mathscr{C}_1}^{\rm OPT}(\text{Medium}) = \{x_2, x_4, x_5, x_8, x_{12}\},\\
&\underline{\mathscr{C}_1}^{\rm OPT}(\text{Good}) = \varnothing;\\
&\underline{\mathscr{C}_2}^{\rm OPT}(\text{Bad}) = \varnothing,\\
&\underline{\mathscr{C}_2}^{\rm OPT}(\text{Medium}) = \{x_5, x_8, x_{12}\},\\
&\underline{\mathscr{C}_2}^{\rm OPT}(\text{Good}) = \varnothing;\\
&\overline{\mathscr{C}_1}^{\rm OPT}(\text{Bad}) = \{x_1, x_3, x_6, x_7, x_{11}, x_{14}\},\\
&\overline{\mathscr{C}_1}^{\rm OPT}(\text{Medium}) = \{x_1, x_2, x_4, x_5, x_8, x_9, x_{10}, x_{12},\\
&\qquad x_{13}, x_{15}\},\\
&\overline{\mathscr{C}_1}^{\rm OPT}(\text{Good}) = \{x_9, x_{10}, x_{13}, x_{15}\};\\
&\overline{\mathscr{C}_2}^{\rm OPT}(\text{Bad}) = \{x_1, x_2, x_3, x_4, x_6, x_7, x_9, x_{10}, x_{11}, x_{13},\\
&\qquad x_{14}, x_{15}\},\\
&\overline{\mathscr{C}_2}^{\rm OPT}(\text{Medium}) = U,\\
&\overline{\mathscr{C}_2}^{\rm OPT}(\text{Good}) = \{x_9, x_{10}, x_{13}, x_{15}\}.
\end{aligned}$$

By the above results, we can see that $\underline{\mathscr{C}_1}^{\rm OPT}(X) \supseteq \underline{\mathscr{C}_2}^{\rm OPT}(X)$ and $\overline{\mathscr{C}_1}^{\rm OPT}(X) \subseteq \overline{\mathscr{C}_2}^{\rm OPT}(X)$, in which X is the decision class in $U/IND(\{d\})$. In other words, this example shows the relationships between the first hierarchical structure and optimistic multicovering rough set in detail.

Theorem 25. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $K(\mathscr{C}_1) \preceq_2 K(\mathscr{C}_2)$, then $\underline{\mathscr{C}_1}^{\rm PES}(X) \supseteq \underline{\mathscr{C}_2}^{\rm PES}(X)$ for each $X \subseteq U$.*

Proof. The proof of Theorem 25 is similar to the proof of Theorem 11. □

Theorem 25 tells us that by using the second hierarchical structure, if the CBMS is finer, then the greater pessimistic multicovering lower approximation will be obtained.

Theorem 26. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $K(\mathscr{C}_1) \preceq_2 K(\mathscr{C}_2)$, then $\overline{\mathscr{C}_1}^{\rm PES}(X) \subseteq \overline{\mathscr{C}_2}^{\rm PES}(X)$ for each $X \subseteq U$.*

Proof. The proof of Theorem 26 is similar to the proof of Theorem 11. □

Theorem 26 tells us that by using the second hierarchical structure, if the CBMS is finer, then the smaller pessimistic multicovering upper approximation will be obtained.

Theorem 27. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $\underline{\mathscr{C}_1}^{\mathrm{PES}}(X) \supseteq \underline{\mathscr{C}_2}^{\mathrm{PES}}(X)$ for each $X \subseteq U$, then $K(\mathscr{C}_1) \preceq_2 K(\mathscr{C}_2)$.*

Proof. The proof of Theorem 27 is similar to the proof of Theorem 13. □

Theorem 27 is the inverse of Theorem 25, i.e., if the pessimistic multicovering lower approximation is greater, then the multigranulation space is finer in terms of the second hierarchical structure.

Theorem 28. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $\overline{\mathscr{C}_1}^{\mathrm{PES}}(X) \subseteq \overline{\mathscr{C}_2}^{\mathrm{PES}}(X)$ for each $X \subseteq U$, then $K(\mathscr{C}_1) \preceq_2 K(\mathscr{C}_2)$.*

Proof. The proof of Theorem 28 is similar to the proof of Theorem 13. □

Theorem 28 is the inverse of Theorem 26, i.e., if the pessimistic multicovering upper approximation is smaller, then the multigranulation space is finer in terms of the second hierarchical structure.

Example 10. Take for instance the table used in Example 9, by Definition 12, we know that $K(\mathscr{C}_1) \preceq_2 K(\mathscr{C}_2)$ holds. Therefore, it is not difficult to obtain the following pessimistic multicovering rough sets:

$$\begin{aligned}
&\underline{\mathscr{C}_1}^{\mathrm{PES}}(\mathrm{Bad}) = \{x_3, x_6, x_7, x_{11}, x_{14}\},\\
&\underline{\mathscr{C}_1}^{\mathrm{PES}}(\mathrm{Medium}) = \{x_2, x_4, x_5, x_8, x_{12}\},\\
&\underline{\mathscr{C}_1}^{\mathrm{PES}}(\mathrm{Good}) = \varnothing;\\
&\underline{\mathscr{C}_2}^{\mathrm{PES}}(\mathrm{Bad}) = \varnothing,\\
&\underline{\mathscr{C}_2}^{\mathrm{PES}}(\mathrm{Medium}) = \varnothing,\\
&\underline{\mathscr{C}_2}^{\mathrm{PES}}(\mathrm{Good}) = \varnothing;\\
&\overline{\mathscr{C}_1}^{\mathrm{PES}}(\mathrm{Bad}) = \{x_1, x_3, x_6, x_7, x_{11}, x_{14}\},\\
&\overline{\mathscr{C}_1}^{\mathrm{PES}}(\mathrm{Medium}) = \{x_1, x_2, x_4, x_5, x_8, x_9, x_{10}, x_{12},\\
&\qquad\qquad x_{13}, x_{15}\},\\
&\overline{\mathscr{C}_1}^{\mathrm{PES}}(\mathrm{Good}) = \{x_9, x_{10}, x_{13}, x_{15}\};\\
&\overline{\mathscr{C}_2}^{\mathrm{PES}}(\mathrm{Bad}) = U,\\
&\overline{\mathscr{C}_2}^{\mathrm{PES}}(\mathrm{Medium}) = U,\\
&\overline{\mathscr{C}_2}^{\mathrm{PES}}(\mathrm{Good}) = U.
\end{aligned}$$

By the above results, we can see that $\underline{\mathscr{C}_1}^{\mathrm{PES}}(X) \supseteq \underline{\mathscr{C}_2}^{\mathrm{PES}}(X)$ and $\overline{\mathscr{C}_1}^{\mathrm{PES}}(X) \subseteq \overline{\mathscr{C}_2}^{\mathrm{PES}}(X)$, in which X is the decision class in $U/IND(\{d\})$. In other words, this example shows the relationships between the second hierarchical structure and pessimistic multicovering rough set in detail.

Since the third hierarchical structure can generate the first and the second hierarchical structures (see Fig.2), then by the above theorems, it is not difficult to obtain the following corollaries.

Corollary 5. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $K(\mathscr{C}_1) \preceq_3 K(\mathscr{C}_2)$, then $\underline{\mathscr{C}_1}^{\mathrm{OPT}}(X) \supseteq \underline{\mathscr{C}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$.*

Corollary 6. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $K(\mathscr{C}_1) \preceq_3 K(\mathscr{C}_2)$, then $\overline{\mathscr{C}_1}^{\mathrm{OPT}}(X) \subseteq \overline{\mathscr{C}_2}^{\mathrm{OPT}}(X)$ for each $X \subseteq U$.*

Corollary 7. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $K(\mathscr{C}_1) \preceq_3 K(\mathscr{C}_2)$, then $\underline{\mathscr{C}_1}^{\mathrm{PES}}(X) \supseteq \underline{\mathscr{C}_2}^{\mathrm{PES}}(X)$ for each $X \subseteq U$.*

Corollary 8. *Let U be the universe of discourse, $\mathscr{C}_1$ and $\mathscr{C}_2$ are two families of the coverings on U, if $K(\mathscr{C}_1) \preceq_3 K(\mathscr{C}_2)$, then $\overline{\mathscr{C}_1}^{\mathrm{PES}}(X) \subseteq \overline{\mathscr{C}_2}^{\mathrm{PES}}(X)$ for each $X \subseteq U$.*

The above four corollaries tell us that if the CBMS is finer, then the optimistic and pessimistic multicovering lower approximations are greater and the optimistic and pessimistic multicovering upper approximations are smaller.

6 Conclusions

In this paper, two different multigranulation spaces, partition-based multigranulation space and covering-based multigranulation space have been deeply investigated. Different from single granulation space, multigranulation space is derived from a family of the binary relations or a family of the coverings. Moreover, three different hierarchical structures are proposed on such two multigranulation spaces, respectively. These hierarchical structures can be used to explore the finer or coarser relationships between different multigranulation spaces.

It is proven that the third hierarchical structure can induce the first and the second hierarchical structures, the first hierarchical structure is corresponding to the monotonic varieties of the optimistic lower and upper approximations, and the second hierarchical structure is corresponding to the monotonic varieties of the pessimistic lower and upper approximations. These results show that the hierarchical structures proposed in this paper may be better for characterizing the essence of multigranulation rough sets, which will be very helpful for establishing a uniform framework for granular computing.

The present study is the first step to the development of multigranulation space and multigranulation rough sets. The following are challenges for further research:

1) The uncertainty measurements are interesting

issues to be addressed in multigranulation spaces, and the relationships between the uncertainty measurements and the proposed hierarchical structures can also be examined.

2) In our paper, the multicovering rough set is generalized by one type of the single granulation covering rough set; since many different types of covering rough sets have been proposed, and then how to generalize these covering rough sets into multigranulation framework and compare the relationships among different multicovering rough sets remain problems for future investigation.

3) Optimism and pessimism are two special approaches for the constructing of multigranulation rough sets, and new models for knowledge acquisition under multigranulation framework is a challenge for further analysis.

References

[1] von Neumann J. The general and logical theory of automata. In *Cerebral Mechanisms in Behavior — The Hixon Symposium*, Jeffress L A (ed.), John Wiley & Sons, 1951, pp.1-41.

[2] Zadeh L. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets and Systems*, 1997, 90(2): 111-127.

[3] Pedrycz W. Relational and directional aspects in the construction of information granules. *IEEE Trans. Systems, Man, and Cybernetics, Part A*: *Systems and Humans*, 2002, 32(5): 605-614.

[4] Pedrycz W, Bargiela A. Granular clustering: a granular signature of data. *IEEE Trans. Systems, Man, and Cybernetics, Part B*: *Cybernetics*, 2002, 32(2): 212-224.

[5] Lin T Y. Granular computing: Practices, theories, and future directions. In *Encyclopedia of Complexity and Systems Science*, Meyers R A (ed.), Springer, 2009, pp.4339-4355.

[6] Lin T Y. Granular computing I: The concept of granulation and its formal model. *Int. J. Granular Computing, Rough Sets and Intell. Systems*, 2009, 1(1): 21-42.

[7] Pawlak Z. Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, 1991.

[8] Zhang L, Zhang B. Theory and Application of Problem Solving: Theory and Application of Granular Computing in Quotient Space. Tsinghua University Press, 2007. (In Chinese)

[9] Yang X B, Zhang M, Dou H L, Yang J Y. Neighborhood systems-based rough sets in incomplete information system. *Knowledge–Based Systems*, 2011, 24(6): 858-867.

[10] Zhu W. Topological approaches to covering rough sets. *Information Sciences*, 2007, 177(6): 1499-1508.

[11] Zhu W. Relationship between generalized rough sets based on binary relation and covering. *Information Sciences*, 2009, 179(3): 210-225.

[12] Zhu W, Wang F Y. On three types of covering-based rough sets. *IEEE Trans. Knowledge and Data Engineering*, 2007, 19(8): 1131-1144.

[13] Qian Y H, Liang J Y. Rough set method based on multigranulations. In *Proc. the 5th IEEE Int. Conf. Cognitive Informatics*, July 2006, pp.297-304.

[14] Qian Y H, Liang J Y, Dang C Y. Incomplete multigranulation rough set. *IEEE Trans. Systems, Man, and Cybernetics, Part A*: *Systems and Humans*, 2010, 40(2): 420-431.

[15] Qian Y H, Liang J Y, Yao Y Y, Dang C Y. MGRS: A multigranulation rough set. *Information Sciences*, 2010, 180(6): 949-970.

[16] Qian Y H, Liang J Y, Wei W. Pessimistic rough decision. *Journal of Zhejiang Ocean University* (*Natural Science*), 2010, 29(5): 440-449.

[17] Yang X B, Song X N, Dou H L, Yang J Y. Multi-granulation rough set: From crisp to fuzzy case. *Annals of Fuzzy Mathematics and Informatics*, 2011, 1(1): 55-70.

[18] Wu W Z, Leung Y. Theory and applications of granular labelled partitions in multi-scale decision tables. *Information Sciences*, 2011, 181(18): 3878-3897.

[19] Qian Y H, Liang J Y, Pedrycz W, Dang C Y. Positive approximation: An accelerator for attribute reduction in rough set theory. *Artificial Intelligence*, 2010, 174(9-10): 597-618.

[20] Yao Y Y. Information granulation and rough set approximation. *Int. J. Intell. Syst.*, 2001, 16(1): 87-104.

[21] Wang G Y, Zhang Q H. Uncertainty of rough sets in different knowledge granularities. *Chinese Journal of Computers*, 2008, 31(9): 1588-1598. (In Chinese)

[22] Huang B, He X, Zhou X Z. Rough entropy based on generalized rough sets covering reduction. *Journal of Software*, 2004, 15(2): 215-220. (In Chinese)

[23] Zhang Q H, Wang G Y, Hu J, Teng H T. Approximation partition spaces of covering space. In *Proc. the 2007 IEEE Int. Conf. Granular Computing*, November 2007, pp.199-204.

[24] Zhang Q H, Wang G Y, Liu X Q. Hierarchical structure analysis of fuzzy quotient space. *Pattern Recognition and Artificial Intelligence*, 2008, 21(5): 627-6434. (In Chinese)

[25] Zhang Q H, Wang G Y. The uncertainty measure of hierarchical quotient space structure. *Mathematical Problems in Engineering*, 2011, Article No. 513195.

[26] Liang J Y, Chin K S, Dang C Y. A new method for measuring uncertainty and fuzziness in rough set theory. *Int. J. General Systems*, 2002, 31(4): 331-342.

[27] Liang J Y, Shi Z Z. The information entropy, rough entropy and knowledge granulation in rough set theory. *Int. J. Uncertainty Fuzziness and Knowledge-Based Systems*, 2004, 12(1): 37-46.

[28] Qian Y H, Liang J Y, Dang C Y. Knowledge structure, knowledge granulation and knowledge distance in a knowledge base. *Int. J. Approximation Reasoning*, 2009, 50(1): 174-188.

[29] Qian Y H, Liang J Y, Wu W Z, Dang C Y. Information granularity in fuzzy binary GrC model. *IEEE Trans. Fuzzy Systems*, 2011, 19(2): 253-264.

[30] Liu G L, Sai Y. A comparison of two types of rough sets induced by coverings. *Int. J. Approximation Reasoning*, 2009, 50(3): 521-528.

[31] Leung Y, Li D Y. Maximal consistent block technique for rule acquisition in incomplete information systems. *Information Sciences*, 2003, 153(1): 85-106.

Xi-Bei Yang received the B.S. degree from Xuzhou Normal University, China, in 2002, the M.S. degree from the Jiangsu University of Science and Technology, China, in 2006, and the Ph.D. degree from Nanjing University of Science and Technology, China, in 2010, both in computer applications. He is currently a lecturer at Jiangsu University of Science and Technology, and also a postdoctoral researcher at Nanjing University of Science and Technology. He has published more than 30 articles in international journals and international conferences. His research interests include granular computing, rough set, and decision making.

Yu-Hua Qian received the M.S. and the Ph.D. degrees in computers with applications from Shanxi University, China, in 2005 and 2011, respectively. He is a professor with the Key Laboratory of Computational Intelligence and Chinese Information Processing of the Ministry of Education, China. He is actively pursuing research in pattern recognition, feature selection, rough set theory, granular computing, and artificial intelligence. He has published more than 40 articles in international journals. Dr. Qian has also served on the editorial board of the International Journal of Knowledge-Based Organizations and that of Artificial Intelligence Research.

Jing-Yu Yang was a visiting scientist at the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, USA, from 1982 to 1984. From 1993 to 1994, he was a visiting professor in the Department of Computer Science, Missuria University, USA. And, in 1998, he acted as a visiting professor at Concordia University in Canada. He is currently a professor in the School of Computer Science and Technology at Nanjing University of Science and Technology, China. He is the author of more than 300 articles. He has won more than 20 provincial and national awards. His current research interests are in the areas of pattern recognition, robot vision, image processing, data fusion, and artificial intelligence.