Granulation Operators on a Knowledge Base

Yuhua Qian Jiye Liang Wei Wei
Key Laboratory of Computational Intelligence and Chinese Information Processing of
Ministry of Education, Taiyuan, 030006, Shanxi, China
School of Computer and Information Technology, Shanxi University, Taiyuan, 030006, Shanxi, China
jinchengqyh@126.com ljy@sxu.edu.cn weiwei@sxu.edu.cn

Abstract

Knowledge in knowledge bases have two categories: complete and incomplete. In this paper, through uniformly expressing these two kinds of knowledge, we first address four operators on a knowledge base, which are adequate for generating new knowledge through using known knowledge. Then, we establish the relationship between knowledge and knowledge granulation. These results will be very helpful for knowledge discovery from knowledge bases and play a significant role for establishing a framework of granular computing in knowledge bases.

Keywords*: Granular computing; rough set theory; knowledge base; granulation operator.*

I. Introduction

Rough set theory, proposed by Pawlak [14, 16], has become a well-established mechanism for uncertainty management in a wide variety of applications related to artificial intelligence [1, 3, 4, 13, 25, 26, 28, 33]. In this framework, an attribute set is viewed as a family of knowledge, which partitions the universe into some knowledge granules or elemental concepts. In other words, instead of using external numbers or other additional parameters, the rough set data analysis (RSDA) utilizes solely the granularity structure of the given data, expressed as classes of suitable equivalence relations. Partition, granulation and approximation are the methods widely used in human's reasoning [17, 31, 32, 34].

Knowledge bases and indiscernibility relations are two basic concepts in the rough set theory and assessing the uncertainty of knowledge in a knowledge base is an important research issue [30]. According to whether or not there missing data (null values), knowledge bases can be classified into two categories: complete and incomplete [5-7]. In the rough set theory, information entropy and knowledge granulation are two main approaches for measuring the uncertainty of a knowledge in knowledge bases [12, 24]. As follows, for our further development, we briefly review several existing knowledge granulations. Wierman [27] presented a well justified measure of uncertainty, the measure of granularity, along with an axiomatic derivation. Its strong connections to the Shannon entropy and the Hartley measure of uncertainty [2] also lend strong support to its correctness and applicability. Liang et al. in the literature [8] and [9] gave the definitions of knowledge granulation in a complete knowledge base and an incomplete knowledge base, respectively, and established the relationships among information entropy, rough entropy and knowledge granulation in knowledge bases. Qian and Liang [20, 24] introduced a new knowledge granulation, called combination granulation, for measuring the uncertainty of a knowledge in complete/incomplete knowledge bases.

For a given knowledge base, one of tasks in data mining and knowledge discovery is to generate new knowledge through using known knowledge. However, the number of knowledge is finite in a given knowledge base, which limits the ability of this knowledge base for approximating an unknown concept. This leads to a task for acquiring more knowledge from a given knowledge base. To date, the mechanism that how to generate new knowledge based on known knowledge in knowledge bases have not been widely researched. Therefore, such a mechanism is desirable and will be very helpful for rule extraction and knowledge discovery from knowledge bases. Based on these analyses, main objective of this study is to establish a new mathematical framework of granular computing in the context of knowledge bases. We focus mainly on granulation operators on knowledge bases in the present research.

The rest of the paper is organized as follows. Some basic concepts in rough set theory are briefly reviewed in Section 2. In Section 3, we establish four operators ($\bigcap, \bigcup, \wr$ and $-$) on a knowledge base and investigate their operation properties. Noting that $(K, \bigcap, \bigcup)$ is an assignment lattice and $(K, \bigcap, \bigcup, \wr)$ is a complemented lattice. Finally, Section 4 concludes this paper with some remarks and discussions.

II. Knowledge in a knowledge base

In this section, we will review several basic concepts in rough set theory and knowledge bases. Throughout this paper, we suppose that the universe U is a finite nonempty set.

Let U be a finite and non-empty set called the universe and $R \subseteq U \times U$ an equivalence relation on U, then $K = (U, R)$ is called an approximation space [14]. The equivalence relation R partitions the set U into disjoint subsets. This partition of the universe is called a quotient set induced by R, denoted by U/R. It represents a very special type of similarity between elements of the universe. If two elements $x, y \in U(x \neq y)$ belong to the same equivalence class, we say that x and y are indistinguishable under the equivalence relation R, i.e., they are equal in R. We denote the equivalence class including x by $E_R(x)$. $K(R)$ is called a knowledge induced by U/R on U. Each equivalence class $E_R(x)(x \in U)$ may be viewed as a knowledge granule consisting of indistinguishable elements [15, 19, 21, 22, 30]. The granulation structure induced by an equivalence relation is a partition of the universe.

We say $K = (U, \mathbf{R})$ is a knowledge base, where U is a finite and non-empty set and $\mathbf{R}$ is a family of equivalence relations. Through using a given knowledge, one can construct a rough set of any subset on the universe in the following definition.

Let $K = (U, \mathbf{R})$ be a knowledge base, if $R(R \in \mathbf{R})$ is an equivalence relation, then we can get a cover of U by $U/R = \{E_R(x) \mid x \in U\}$, i.e., for $\forall x \in U$, one has that $E_R(x) \neq \emptyset$ and $\bigcup_{x \in U} E_R(x) = U$. Obviously, $\forall x, y \in U(x \neq y)$, if x, y are partitioned into the same equivalence class, then $E_R(x) = E_R(y)$, otherwise $E_R(x) \cap E_R(y) = \emptyset$. One can define a partial relation $\preceq$ as follows: $P \preceq Q$ $(P, Q \in \mathbf{R})$ if and only if, one has $E_P(x_i) \subseteq E_Q(x_i)$ for any $i \in \{1, 2, \cdots, |U|\}$ [9-11, 18, 20, 23]. Here, we denote that P is finer than Q by $P \preceq Q$. Obviously, $(\mathbf{R}, \preceq)$ is a poset [29].

Similarly, let $R \subseteq U \times U$ denote a tolerance relation on U, the tolerance relation R classifies the universe U into some subsets, i.e., a cover of U [5, 6]. This cover of the universe is called a knowledge induced by R, denoted by U/R or $K(R)$. If y belongs to the tolerance class determined by x with respect to R, we say two elements x, y are indistinguishable under the tolerance relation R, i.e., they are similar in R [5-7]. We denote the tolerance class of x by $S_R(x)$ [5, 10, 11]. Each tolerance class $S_R(x)(x \in R)$ is viewed as a knowledge granule [9, 18, 20]. The granulation structure induced by a tolerance relation is a cover of the universe. Conveniently, we say $K = (U, \mathbf{R})$ is also a knowledge base, where U is a finite and non-empty set and $\mathbf{R}$ is a family of tolerance relations. The following definition gives a rough set of a subset of the universe based on a tolerance relation.

Let $K = (U, \mathbf{R})$ be a knowledge base, if $R(R \in \mathbf{R})$ is a tolerance relation, then we can denote a cover of U by $U/R = \{S_R(x) \mid x \in U\}$, i.e., $\forall x \in U$, one has $S_R(x) \neq \emptyset$ and $\bigcup_{x \in U} S_R(x) = U$. In [9, 10], Liang et al. defined a partial relation $\preceq$ as follows: $P \preceq Q (P, Q \in \mathbf{R})$ if and only if, for every $i \in \{1, 2, \cdots, |U|\}$, one has that $S_P(x_i) \subseteq S_Q(x_i)$. Here, we also denote that P is finer than Q by $P \preceq Q$. It is easy to see that $(\mathbf{R}, \preceq)$ is also a poset.

III. Granulation operators with properties

In this section, by uniformly representing a complete knowledge and an incomplete knowledge, we will propose four granulation operators on a knowledge base and discuss their fundamental algebra properties.

In [8, 9], liang et al. established the relationship between a complete knowledge and an incomplete knowledge in the same knowledge base. Let $K = (U, R)$ be a knowledge, R a equivalence relation, $U/R = \{X_1, X_2, \cdots, X_m\}$, $U/R = \{S_R(x_1), S_R(x_2), \cdots, S_R(x_{|U|})\}$ and $X_i = \{x_{i1}, x_{i2}, \cdots, x_{is_i}\}$, where $|X_i| = s_i$ and $\sum_{i=1}^{m} s_i = |U|$, then

$$X_i = S_R(x_{i1}) = S_R(x_{i2}) = \cdots = S_R(x_{is_i}). \quad (1)$$

Through this mechanism, one can denote $U/R = \{E_R(x) \mid x \in U\}$ by using $U/R = \{S_R(x) \mid x \in U\}$. The mechanism gives uniform representations of knowledge in a knowledge base. It is illustrated by the following example.

Example 1: Let $U = \{x_1, x_2, \cdots, x_6\}$, R a equivalence relation and $U/R = \{\{x_1, x_2\}, \{x_3, x_4, x_5\}, \{x_6\}\}$. Then, $U/R = \{S_R(x) \mid x \in U\}$ can be represented equivalently as

$$\begin{aligned} U/R &= \{S_R(x_1), S_R(x_2), S_R(x_3), S_R(x_4), S_R(x_5), S_R(x_6)\} \\ &= \{\{x_1, x_2\}, \{x_1, x_2\}, \{x_3, x_4, x_5\}, \{x_3, x_4, x_5\}, \\ &\quad \{x_3, x_4, x_5\}, \{x_6\}\}. \quad \square \end{aligned}$$

For convenience, we denote the knowledge induced by R on U as $K(R)$ in the rest of this paper, where R is an equivalence relation or a tolerance relation.

There are two types of operators to be considered in granular computing based on rough set theory. One is operations among knowledge granules, the other is operations among knowledge in a knowledge base. As operations among knowledge granules is based on classical sets, we still operate on them by $\cap$, $\cup$, $-$ and $\sim$, i.e., a new knowledge granule can be generated by $\cap$, $\cup$, $-$ and $\sim$ on known knowledge granules. However, operations among knowledge are performed through composing and decomposing unknown knowledge in knowledge bases in essence. Therefore, the operators on a knowledge base to generate new knowledge are very desirable. In the following, we introduce four granulation operators among knowledge in a knowledge base.

Definition 1: Let $K = (U, \mathbf{R})$ be a knowledge base and $K(P)$, $K(Q) \in K$ two knowledge. Four operators $\bigcap$, $\bigcup$, $-$ and $\wr$ on K are defined as

$$K(P) \bigcap K(Q) = \{S_{P\cap Q}(x) \mid S_{P\cap Q}(x) = S_P(x) \cap S_Q(x),\ x \in U\},$$

$$K(P) \bigcup K(Q) = \{S_{P\cup Q}(x) \mid S_{P\cup Q}(x) = S_P(x) \cup S_Q(x),\ x \in U\},$$

$$K(P) - K(Q) = \{S_{P-Q}(x) \mid S_{P-Q}(x) = x \cup (S_P(x) - S_Q(x)), x \in U\},$$

and

$$\wr K(P) = \{\wr S_P(x) \mid \wr S_P(x) = x \cup \sim S_P(x), x \in U\},$$

where $\sim S_P(x) = U - S_P(x)$.

Here, we regard $\bigcap$, $\bigcup$ $-$ and $\wr$ as four atomic formulas and finite connection on them are all formulas. Through using these operators, one can obtain new knowledge via some known knowledge on U. Let $\mathbf{K}(U)$ denote the set of all knowledge on U, then these four operators $\bigcap$, $\bigcup$, $-$ and $\wr$ on $\mathbf{K}(U)$ are close. As follows, we investigate several fundamental algebra properties of these four operators.

Theorem 1: Let $\bigcap$, $\bigcup$ be two operators on K, then

(1) $K(P) \bigcap K(P) = K(P)$,
$K(P) \bigcup K(P) = K(P)$;

(2) $K(P) \bigcap K(Q) = K(Q) \bigcap K(P)$,
$K(P) \bigcup K(Q) = K(Q) \bigcup K(P)$;

(3) $K(P) \bigcap (K(P) \bigcup K(Q)) = K(P)$,
$K(P) \bigcup (K(P) \bigcap K(Q)) = K(P)$; and

(4)
$(K(P) \bigcap K(Q)) \bigcap K(R) = K(P) \bigcap (K(Q) \bigcap K(R))$,
$(K(P) \bigcup K(Q)) \bigcup K(R) = K(P) \bigcup (K(Q) \bigcup K(R))$.

Proof. They are straightforward from definition 3. $\square$

Theorem 2: Let $\bigcap$, $\bigcup$ and $\wr$ be three operators on K. Then,

(1) $\wr(\wr K(P)) = K(P)$,

(2) $K(P) \bigcap \wr K(P) = \{x_i \mid x_i \in U\}$,

(3) $\wr(K(P) \bigcap K(Q)) = \wr K(P) \bigcup \wr K(Q)$, and

(4) $\wr(K(P) \bigcup K(Q)) = \wr K(P) \bigcap \wr K(Q)$.

Proof. Let $K(P), K(Q) \in K$. For any $x_i \in U$, $S_P(x_i)$ is the tolerance class induced by x_i in $K(P)$ and $S_Q(x_i)$ is the tolerance class induced by x_i in $K(Q)$.

(1) From Definition 13, one can easily see that $\wr(S_P(x_i)) = x_i \cup \sim S_P(x_i)$ and $\wr(\wr(S_P(x_i))) = x_i \cup (x_i \cup S_P(x_i)) = S_P(x_i)$. Therefore, $\wr(\wr K(P)) = K(P)$.

(2) From Definition 1, it follows that $S_P(x_i) \cap \wr(S_P(x_i)) = x_i$, $\forall x_i \in U$. Then, $K(P) \bigcap \sim K(P) = \{x_i \mid x_i \in U\}$.

(3) According to Definition 1, for $\forall x_i \in U$, it follows that

$$\begin{aligned}\wr(S_P(x_i) \cap S_Q(x_i)) &= x_i \cup \sim (S_P(x_i) \cap S_Q(x_i)) \\ &= x_i \cup (\sim S_P(x_i) \cup \sim S_Q(x_i)) \\ &= (x_i \cup \sim S_P(x_i)) \cup (x_i \cup \sim S_Q(x_i)) \\ &= \wr S_P(x_i) \cup \wr S_Q(x_i).\end{aligned}$$

Therefore, one can get that

$$\wr(K(P) \bigcap K(Q)) = \wr K(P) \bigcup \wr K(Q).$$

(4) According to Definition 1, for $\forall x_i \in U$, one has that

$$\begin{aligned}\wr(S_P(x_i) \cup S_Q(x_i)) &= x_i \cup \sim (S_P(x_i) \cup S_Q(x_i)) \\ &= x_i \cup (\sim S_P(x_i) \cap \sim S_Q(x_i)) \\ &= (x_i \cup \sim S_P(x_i)) \cap (x_i \cup \sim S_Q(x_i)) \\ &= \wr S_P(x_i) \cap \wr S_Q(x_i).\end{aligned}$$

Hence, one can obtain that

$$\wr(K(P) \bigcup K(Q)) = \wr K(P) \bigcap \wr K(Q). \quad \square$$

Theorem 2 shows that (1) is reflexive, (2) is complementary, and (3) and (4) are two dual principles.

Theorem 3: Let $\bigcap$, $\bigcup$, $-$ and $\wr$ be operators on K. Then,

(1) $K(P) - K(Q) = K(P) \bigcap \wr K(Q)$,

(2) $K(P) - K(Q) = K(P) - (K(P) \bigcap K(Q))$,

(3) $K(P) \bigcap (K(Q) - K(R)) = (K(P) \bigcap K(Q)) - (K(P) \bigcap K(R))$, and

(4) $(K(P) - K(Q)) \bigcup K(Q) = K(P)$.

Proof. They are straightforward from Definition 1. $\square$

The above three theorems are illustrated by the following example.

Example 2: Let $U = \{x_1, x_2, x_3, x_4\}$, $K(P) = \{\{x_1,\ x_2\},\ \{x_1, x_2\}, \{x_3, x_4\},\ \{x_3,\ x_4\}\}$ and $K(Q) = \{\{x_1, x_4\},\ \{x_2, x_3\}, \{x_2, x_3, x_4\}, \{x_4\}\}$, one can acquire some new knowledge through using $K(P)$ and $K(Q)$.

By computing, some new knowledge constructed are listed as follows.

$\wr K(P) = \{\{x_1, x_3, x_4\}, \{x_2, x_3, x_4\}, \{x_1, x_2, x_3\}, \{x_1, x_2, x_4\}\}$,

$\wr K(Q) = \{\{x_1, x_2, x_3\}, \{x_1, x_2, x_4\}, \{x_1, x_3\}, \{x_1, x_2, x_3, x_4\}\}$,

$K(P) \bigcap K(Q) = \{\{x_1\}, \{x_2\}, \{x_3, x_4\}, \{x_4\}\}$,

$K(P) \bigcup K(Q) = \{\{x_1, x_2, x_4\}, \{x_1, x_2, x_3\}, \{x_2, x_3, x_4\}, \{x_1, x_4\}\}$,

$\wr K(P) \bigcap \wr K(Q) = \{\{x_1, x_3\}, \{x_2, x_4\}, \{x_1, x_3\}, \{x_1, x_2, x_4\}\}$,

$\wr K(P) \bigcup \wr K(Q) = \{\{x_1, x_2, x_3, x_4\}, \{x_1, x_2, x_3, x_4\}, \{x_1,$

$x_2, x_3\}, \{x_1, x_2, x_3, x_4\}\}$,
$K(P) - K(Q) = \{\{x_1, x_2\}, \{x_1, x_2\}, \{x_3\}, \{x_3, x_4\}\}$ and $\wr K(P) - \wr K(Q) = \{\{x_1, x_4\}, \{x_2, x_3\}, \{x_2, x_3\}, \{x_4\}\}$. □

Suppose $K = (U, R)$ be a knowledge base, $P, Q \in R$, and $K(P)$, $K(Q) \in K$ be two knowledge induced by P, Q, respectively. To investigate properties of the operations among knowledge on a knowledge base, we will write $K(P) \preceq K(Q)$ iff $P \preceq Q$.

Theorem 4: Let $\bigcap$, $\bigcup$ and $\wr$ be three operators on K. The following properties hold:

(1) If $K(P) \preceq K(Q)$, then $\wr K(Q) \preceq \wr K(P)$;

(2) $K(P) \bigcap K(Q) \preceq K(P)$, $K(P) \bigcap K(Q) \preceq K(Q)$; and

(3) $K(P) \preceq K(P) \bigcup K(Q)$, $K(Q) \preceq K(P) \bigcup K(Q)$.

Proof. The terms (2) and (3) can be easily proved from Definition 1, respectively.

From Definition 1, one can obtain that

$K(P) \preceq K(Q)$
$\Longrightarrow$ for $\forall x_i \in U$, $S_P(x_i) \subseteq S_Q(x_i)$
$\Longrightarrow$ for $\forall x_i \in U$, $\sim S_Q(x_i) \subseteq \sim S_P(x_i)$
$\Longrightarrow$ for $\forall x_i \in U$, $x_i \cup \sim S_Q(x_i) \subseteq x_i \cup \sim S_P(x_i)$
$\Longrightarrow \wr K(Q) \preceq \wr K(P)$.

Hence, the term (1) in this theorem holds. □

Definition 2: [35] Let $(L, \leq)$ be a poset. If there exist two operators $\wedge, \vee$ on $L : L^2 \longrightarrow L$ such that

(1) $a \wedge b = b \wedge a$, $a \vee b = b \vee a$,

(2) $(a \wedge b) \wedge c = a \wedge (b \wedge c)$, $(a \vee b) \vee c = a \vee (b \vee c)$, and

(3) $a \wedge b = b \Longleftrightarrow b \leq a$, $a \vee b = b \Longleftrightarrow a \leq b$,

then we call L is a lattice.

Furthermore, if

(4) $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ and $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$,

then we call L is an assignment lattice.

We call L a complemented lattice, if for any $a \in L$, there exists a' such that $(a')' = a$ and $a \leq b \Longleftrightarrow b' \leq a'$. If there exist $0, 1 \in L$ such that $0 \leq a \leq 1$ for any $a \in L$, then we call 0 and 1 its minimal element and maximal element, respectively.

Theorem 5: $(K, \bigcup, \bigcap)$ is an assignment lattice.

Proof. At first, we prove $(K, \preceq)$ is a lattice.

From (2) and (4) in Theorem 1, the terms (1) and (2) in Definition 2 are obvious.

Let $K(P), K(Q), K(R) \in K$ be three knowledge, where $K(P) = \{S_P(x) \mid x \in U\}$, $K(Q) = \{S_Q(x) \mid x \in U\}$ and $K(R) = \{S_R(x) \mid x \in U\}$. one can obtain that

$K(P) \bigcap K(Q) = K(P)$
$\Longleftrightarrow$ for $\forall x_i \in U$, $S_{P \cap Q}(x_i) = S_P(x_i)$, $i \leq |U|$
$\Longleftrightarrow S_P(x_i) \cap S_Q(x_i) = S_P(x_i)$
$\Longleftrightarrow S_P(x_i) \subseteq S_Q(x_i)$, for $\forall x_i \in U$
$\Longleftrightarrow K(P) \preceq K(Q)$.

According to the dual principle in a lattice, one can easily get that $K(P) \bigcup K(Q) = K(P) \Longleftrightarrow K(Q) \preceq K(P)$. Thus, the term (3) in Definition 2 holds.

In addition, for $K(P), K(Q), K(R) \in K$, we know that

$S_P(x_i) \cap (S_Q(x_i) \cup S_R(x_i))$
$= (S_P(x_i) \cap (S_Q(x_i)) \cup (S_P(x_i) \cap S_R(x_i))$, $\forall x_i \in U$.

Hence,

$K(P) \bigcap (K(Q) \bigcup K(R))$
$= (K(P) \bigcap K(Q)) \bigcup (K(P) \bigcap K(R))$.

From the dual principle in a lattice, one can get that

$K(P) \bigcup (K(Q) \bigcap K(R))$
$= (K(P) \bigcup K(Q)) \bigcap (K(P) \bigcup K(R))$.

Therefore, $(K, \bigcup, \bigcap)$ is an assignment lattice. □

Theorem 6: Let $\mathbf{K}(U)$ be the set of all knowledge on U. Then, $(\mathbf{K}(U), \bigcup, \bigcap, \wr)$ is a complemented lattice.

Proof. From the above Theorem 5, it is obvious that $(\mathbf{K}(U), \bigcup, \bigcap, \wr)$ is an assignment lattice. Furthermore, from (1) in Theorem 2, one can get that $\wr(\wr K(P)) = K(P)$. In addition, from (3) in Definition 1, one has that

$K(P) \preceq K(Q)$
$\Longleftrightarrow$ for $\forall x_i \in U$, $S_P(x_i) \subseteq S_Q(x_i)$
$\Longleftrightarrow$ for $\forall x_i \in U$, $\sim S_P(x_i) \supseteq \sim S_Q(x_i)$
$\Longleftrightarrow$ for $\forall x_i \in U$, $x_i \cup \sim S_P(x_i) \supseteq x_i \cup \sim S_Q(x_i)$
$\Longleftrightarrow$ for $\forall x_i \in U$, $\wr S_P(x_i) \supseteq \wr S_Q(x_i)$
$\Longleftrightarrow K(Q) \preceq K(P)$.

Hence, $(\mathbf{K}(U), \bigcup, \bigcap, \wr)$ is a complemented lattice. □

In a complemented lattice $(\mathbf{K}(U), \bigcup, \bigcap, \wr)$, the knowledge $K(\omega) = \{x_i \mid x_i \in U\}$ and the knowledge $K(\delta) = \{S_P(x_i) \mid S_P(x_i) = U, x_i \in U\}$ are two special knowledge, where $K(\omega)$ is the discrete classification and $K(\delta)$ is the indiscrete classification. For any $K(P) \in \mathbf{K}(U)$, one has that $K(\omega) \preceq K(P) \preceq K(\delta)$. Then, we can call $K(\omega)$ and $K(\delta)$ the minimal element and the maximal element on the lattice $(\mathbf{K}(U), \bigcup, \bigcap, \wr)$, respectively.

In the following, we establish the relationship between knowledge and knowledge granulation. The following definition of knowledge granulation was proposed for measuring the uncertainty of knowledge in the context of incomplete knowledge bases.

Definition 3: [9] Let $K = (U, \mathbf{R})$ be a knowledge base, $P \in \mathbf{R}$ and $K(P) = \{S_P(x_1), S_P(x_2), \cdots, S_P(x_{|U|})\}$. Knowledge granulation of the knowledge $K(P)$ is defined as

$$G(P) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|S_P(x_i)|}{|U|}, \quad (2)$$

where $\frac{|S_P(x_i)|}{|U|}$ is the probability of tolerance class $S_P(x_i)$ within the universe U.

If $K(P)=K(\omega)$, $G(P)$ achieves its minimum value $G(P)=\frac{1}{|U|}$; if $K(P)=K(\delta)$, $G(P)$ achieves its maximum value $G(P)=1$. It is obvious that $\frac{1}{|U|}\leq G(P)\leq 1$.

Theorem 7: Let $K=(U,\mathbf{R})$ be a knowledge base, $P\in\mathbf{R}$, $K(P)=\{S_P(x_1),\, S_P(x_2),\, \cdots, S_P(x_{|U|})\}$ and $\wr P$ the relation induced by $\wr K(P)$. Then, $G(P)+G(\wr P)=1+\frac{1}{|U|}$.

Proof. From Definition 3, it follows that

$$\begin{aligned} & G(P)+G(\wr P) \\ &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(x_i)|}{|U|}+\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|\wr S_P(x_i)|}{|U|} \\ &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(x_i)|}{|U|}+\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|x_i\cup\sim S_P(x_i)|}{|U|} \\ &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|U|+1}{|U|} \\ &= 1+\frac{1}{|U|}. \end{aligned}$$

That is, $G(P)+G(\wr P)=1+\frac{1}{|U|}$. □

Theorem 8: Let $\mathbf{K}(U)$ be the set of all knowledge on U and $K(P),K(Q)\in\mathbf{K}(U)$ two knowledge. Then, $G(P)-G(Q)=G(\wr(Q)-G(\wr P)$.

Proof. Obviously, we have that

$$\begin{aligned} & G(\wr P)-G(\wr Q) \\ &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{1+|U|-|S_P(x_i)|}{|U|}-\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{1+|U|-|S_Q(x_i)|}{|U|} \\ &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_Q(x_i)|-|S_P(x_i)|}{|U|} \\ &= -(G(P)-G(Q)), \end{aligned}$$

i.e., $G(P)-G(Q)=G(\wr Q)-G(\wr P)$. □

Remark. One of the strengths of rough set theory is the fact that an unknown target concept can be characterized approximately by existing knowledge in a knowledge base. From the above analyses, it is shown that these four operators ($\bigcup$, $\bigcap$, $\wr$ and $-$) can be applied to generate new knowledge on a knowledge base. That is to say, one can use these new knowledge to approximate an unknown target. Therefore, this mechanism may be used to rule extraction and knowledge discovery from knowledge bases.

IV. Conclusions

One of the strengths of rough set theory is the fact that an unknown target concept can be characterized approximately by existing knowledge in a knowledge base. However, the number of knowledge is very finite in a given knowledge base, which limits the ability of this knowledge base for approximating an unknown concept or decision. In other words, it is very desirable that one acquire more knowledge from a given knowledge base. In this paper, by uniformly representing a complete knowledge and an incomplete knowledge, we have proposed four granulation operators ($\bigcup$, $\bigcap$, $\wr$ and $-$) on a knowledge base, which can be applied to generate a new knowledge or a granulation space. For these four operators, some of their important properties have also been obtained. In particular, $(K,\bigcap,\bigcup)$ is an assignment lattice and $(K,\bigcap,\bigcup,\wr)$ is a complemented lattice. Moreover, we have established the relationship between knowledge and knowledge granulation in a given knowledge base. These results have been shown to be very helpful for knowledge discovery from knowledge bases and play a significant role for establishing a framework of granular computing in knowledge bases.

Acknowledgement

This work was partially supported by the national natural science foundation of China (No. 60773133, No. 70471003, No. 60573074), the high technology research and development program (No. 2007AA01Z165), the foundation of doctoral program research of the ministry of education of China (No. 20050108004) and key project of science and technology research of the ministry of education of China.

References

[1] I. Düntsch and G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence, 106 (1998) 109-137.
[2] R. V. L. Hartley, Transmission of information, The Bell Systems Technical Journal, 7 (1928) 535-563.
[3] Q. Hu, Z. Xie, D. Yu, Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation, Pattern Recognition, 40 (2007) 3509-3521.
[4] R. Jensen, Q. Shen, Fuzzy-rough sets assisted attribute selection, IEEE Transactions on Fuzzy Systems, 15(1) (2007) 73-89.
[5] M. Kryszkiewicz, Rough set approach to incomplete information systems, Information Sciences, 112 (1998) 39-49.
[6] M. Kryszkiewicz, Rules in incomplete information systems, Information Sciences, 113 (1999) 271-292.
[7] Y. Leung, D. Y. Li, Maximal consistent block technique for rule acquisition in incomplete information systems, Information Sciences, 153 (2003) 85-106.
[8] J. Y. Liang, Z. Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 12(1) (2004) 37-46.
[9] J. Y. Liang, Z. Z. Shi, D. Y. Li, M. J. Wierman, The information entropy, rough entropy and knowledge granulation in incomplete information sysetms, International Journal of General Systems, 35(6) (2006) 641-654.
[10] J. Y. Liang, D. Y. Li, Uncertainty and Knowledge Acquisition in Information Systems, Science Press, Beijing, China, 2005.
[11] J. Y. Liang, Z. B. Xu, The algorithm on knowledge reduction in incomplete information systems, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 24(1) (2002) 95-103.
[12] T. Y. Lin, Introduction to special issues on data mining and granular computing, International Journal of Approximate Reasoning, 40 (2005) 1-2.
[13] J. S. Mi, Y. Leung, W. Z. Wu, An uncertainty measure in partition-based fuzzy rough sets, International Journal General Systems, 34 (2005) 77-90.
[14] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Boston, 1991.

[15] Z. Pawlak, Graularity of knowledge, indiscernibility and rough sets, Proceedings of 1998 IEEE International Conference on Fuzzy Systems, (1998) 106-110.
[16] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences, 177 (2007) 3-27.
[17] W. Pedrycz, A. Bargiela, Granular clustering: a granular signature of data, IEEE Transactions on Systems, Man and Cybernetics Part B, 32(2) (2002) 212-224.
[18] Y. H. Qian, C. Y. Dang, J. Y. Liang, MGRS in incomplete information systems, Proceedings of 2007 IEEE International Conference on Granular Computing, (2007) 163-168
[19] Y. H. Qian, J. Y. Liang, Rough set method based on multi-granulations, Proceedings of 5th IEEE Conference on Cognitive Informatics, I (2006) 297-304.
[20] Y. H. Qian, J. Y. Liang, Combination entropy and combination granulation in incomplete information systems, Lecture Notes in Artificial Intelligence, 4062 (2006) 184-190.
[21] Y. H. Qian, J. Y. Liang, C. Y. Dang, Converse approximation and rule extracting from decision tables in rough set theory, Computers and Mathematics with Applications, 55 (2008) 1754-1765.
[22] Y. H. Qian, J. Y. Liang, D. Y. Li, H. Y. Zhang, C. Y. Dang, Measures for evaluating the decision performance of a decision table in rough set theory, Information Sciences, 178 (2008) 181-202.
[23] Y. H. Qian, J. Y. Liang, X. Zhang, C.Y. Dang, Rough set approach under dynamic granulation in incomplete information systems, Lecture Notes in Artificial Intelligence, 4827 (2007) 1-8.
[24] Y. H. Qian, J. Y. Liang, Combination entropy and combination granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 16(2) (2008) 179-193.
[25] M. Rebolledo, Rough intervals-enhancing intervals f or qualitative modeling of technical systems, Artificial Intelligence, 170 (2006) 667-685.
[26] Q. Shen, A. Chouchoulas, A rough-fuzzy approach for generating classification rules, Pattern Recognition, 35(11) (2002) 2425-2438.
[27] M. J. Wierman, Measuring uncertainty in rough set theory, International Journal of General Systems, 28(4) (1999) 283-297.
[28] W. Z. Wu, J. S. Mi, W. X. Zhang, Generalized fuzzy rough sets, Information Sciences, 151 (2003) 263-282.
[29] Y. Y. Yao, Information granulation and rough set approximation, Internatinal Journal of Intelligent Systems, 16(1) (2001) 87-104.
[30] Y. Y. Yao, Perspectives of Granular Computing, Proceedings of 2005 IEEE International Conference on Granular Computing, I (2005) 85-90.
[31] L. Zadeh, Fuzzy logic equals computing with words, IEEE Transactions on Fuzzy Systems, 4(2) (1996) 103-111.
[32] L. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems, 90 (1997) 111-127.
[33] A. Zeng, D. Pan, Q. L. Zheng, H. Peng, Knowledge acquisition based on rough set theory and principal component analysis, IEEE Intelligent Systems, (2006) 78-85.
[34] L. Zhang, B. Zhang, Theory of fuzzy quotient space (methods of fuzzy granular computing), Journal of Software (in Chinese), 14(4) (2003) 770-776.
[35] W. X. Zhang, Y. Leung, W. Z. Wu, Information Systems and Knowledge Acquisition, Science Press, Beijing, China, 2005.