Incomplete Multi-Granulations Rough Set

Yuhua Qian, Jiye Liang and Chuangyin Dang, *Senior Member*, *IEEE*

Abstract—Original rough set model is concerned primarily with the approximation of sets described by a single binary relation on the universe. In the view of granular computing, the classical rough set theory is researched by a single granulation. This article first extends the rough set model based on tolerance relation to an incomplete rough set model based on multi-granulations, where set approximations are defined through using multi tolerance relations on the universe. Then, several elementary measures are proposed for this rough set framework. Finally, a concept of approximation reduct is introduced to characterize the smallest attribute subset that preserves the lower approximation and upper approximation of all decision classes in this rough set model, and several key algorithms are also designed for applying this theory in practical issues.

Index Terms—Granular computing, rough set, information systems, attribute reduction.

1 INTRODUCTION

Rough set theory, proposed by Pawlak [24], [26], has become a well-established mechanism for uncertainty management in a wide variety of applications related to artificial intelligence [4], [11], [12], [17], [43], [44], [58]. One of the strengths of rough set theory is the fact that all its parameters are obtained from the given data. This can be seen in the paragraph from [24]: “The numerical value of imprecision is not pre-assumed, as it is in probability theory or fuzzy sets–but is calculated on the basis of approximations which are the fundamental concepts used to express imprecision of knowledge”. In this framework, an attribute set is viewed as a family of knowledge, which partitions the universe into some knowledge granules or elemental concepts. Partition, granulation and approximation are the methods widely used in human’s reasoning [55], [56]. Rough set methodology presents a novel paradigm to deal with uncertainty and has been applied to feature selection [18], [48], [49], knowledge reduction [9], [16], [51], rule extraction [1], [8], [35], [50], [59], uncertainty reasoning [20], [25], [31], decision evaluation [36], [37], [38] and granular computing [2], [3], [19], [21], [29], [30], [52], [57].

Knowledge representation in the rough set model is done via *information systems* (IS) which are a tabular form of an OBJECT→ ATTRIBUTE VALUE relationship, similar to relational databases. An information system is an ordered triplet $S = (U, AT, f)$, where U is a finite nonempty set of objects, AT is a finite nonempty set of attributes(predictor features), and $f_a : U \to V_a$ for any $a \in AT$ with V_a being the domain of an attribute a. For any $x \in U$, an information vector of x is given by $Inf(x) = (a, f_a(x)) \mid a \in AT$. In particular, a target information system is given by $S = (U, AT, f, D, g)$, where D is a finite nonempty set of decision attributes and $g_d : U \to V_d$ for any $d \in D$ with V_d being the domain of a decision attribute d.

In the past ten years, several extensions of the rough set model have been proposed in terms of various requirements [26], [27], [28], such as variable precision rough set (VPRS) model [60], rough set model based on tolerance relation [13], [14], [32], [45], [47], Bayesian rough set model [46], fuzzy rough set model [5], rough fuzzy set model [5] and fuzzy probabilistic rough set model [10]. In the view of granular computing (proposed by Zadeh [56]), in these models, a target concept is always characterized via the so-called upper and lower approximations under a single granulation, i.e., the concept is depicted by a known knowledge induced by a single relation (such as equivalence relation, tolerance relation and reflexive relation) on the universe. However, this approach to describing a target concept is mainly based on the following assumption:

If P and Q are two sets of predictor features and $X \subseteq U$ is a target concept, then the rough set of X is derived by the quotient set $U/(P \cup Q)$. In fact, the quotient set is equivalent to the formula

$$\widehat{P \cup Q} = \{P_i \cap Q_j : P_i \in U/P, Q_j \in U/Q\}.$$

It implies three strategies:

(1) any two attributes must be independent in information systems,

(2) an intersection operation between any P_i and Q_j can be performed, and

(3) the target concept is approximately described by using the quotient set $U/(P \cup Q)$.

In fact, the target concept is described by using a finer granulation (partitions) formed through combining two known granulations (partitions) induced by

---

This is an extended version of the paper presented at the 2007 IEEE International Conference on Granular Computing, California, USA.

- *Y.H. Qian and J.Y. Liang (Corresponding author) are with the School of Computer and Information Technology, Shanxi University, Taiyuan 030006, Shanxi Province, China (e-mail: jinchengqyh@126.com; ljy@sxu.edu.cn).*
- *C.Y. Dang is with the Department of Manufacture Engineering and Engineering Management, City University of Hong Kong, Hong Kong (e-mail: mecdang@cityu.edu.hk)*

two attribute subsets. Although it generates much finer granulation and more knowledge, the combination/fining destroys the original granulation structure/partitions [34].

However, this assumption cannot be always satisfied or required in many practical issues in general. For example, in some decision making processes, for the same object (or a sample, project and element), there is a contradiction/unconsistent relationship between its values under one attribute set P and those under another attribute set Q. In other words, we can not perform intersection operations between their quotient sets and the target concept can not be approximated by the quotient set $U/(P \cup Q)$. In this case, we often need to describe concurrently the target concept through multi binary relations (e.g. equivalence relation, tolerance relation, reflexive relations and neighborhood relation) on the universe according to user requirements or targets of problem solving [34].

In the view of granular computing, an equivalence relation (or a tolerance relation) on the universe can be regarded as a granulation, and a partition (or a cover) on the universe can be regarded as a granulation space [19], [53], [54]. Hence, the classical rough set theory is based on a single granulation (only one equivalence relation). Note that any attribute set can induce a certain equivalence relation in a complete information system. In the literature [34], to more widely apply the rough set theory in practical applications, Qian and Liang extended Pawlak's single-granulation rough set model to a *multi-granulations rough set model* (MGRS), where the set approximations are defined by multi equivalence relations on the universe. In the literature [39], [40], [41], [42], authors also investigated the approaches to approximation based on many indiscernibility relations for rough approximations. However, in essence, the approximations in these approaches are still based on a singleton granulation induced by an indiscernibility relation, which can be used to knowledge representation in distributive systems and groups of intelligent agents. Furthermore, in the literature [33], Qian et al. proposed several basic views for establishing multi-granulations rough set model in incomplete information systems.

The main objective of this paper is to fully establish a rough set model based on multi tolerance relations in incomplete information systems. The rest of the paper is organized as follows. Some basic concepts in complete MGRS are briefly reviewed in Section 2. In Section 3, a rough set model based on multi tolerance relations, called incomplete MGRS, is proposed in incomplete information systems and some of its important properties are investigated. Several elementary measures for this rough set model are presented as well, which are accuracy measure, quality of approximation and precision of approximation. In Section 4, a concept of approximation reduct is first introduced to the incomplete MGRS, which is based on the so-called upper approximation reduct and lower approximation reduct. Then, two algorithms for computing the upper/under approximation reducts are designed for applying this theory in practical issues. In Section 5, we provide an illustrative example show the actual applicability of the approach proposed. Finally, Section 6 concludes the whole paper.

2 PRELIMINARIES

Throughout this paper, we suppose that the universe U is a finite nonempty set.

Let us first recall a few facts about partitions and equivalence relations. Suppose that $U/IND(P)$ is a partition of U induced by the attribute set P in an information system. If $x \in U$, let $[x]_P$ be the class containing x in $U/IND(P)$ and θ_P be the equivalence relation associated with $U/IND(P)$, i.e., $x\theta_P y \Leftrightarrow [x]_P = [y]_P$.

Then, we review the set approximations in MGRS. In this rough set model, unlike Pawlak's rough set theory, a target concept is approximated through multi partitions induced by multi equivalence relations [34]. Suppose that $S = (U, AT, f)$ be a complete information system, $X \subseteq U$ and $P_1, P_2, \cdots, P_m$ be m attribute subsets. We define a lower approximation and a upper approximation of X related to $P_1, P_2, \cdots, P_m$ by the following

$$\underline{\sum_{i=1}^{m} P_i} X = \bigcup\{x \mid [x]_{P_i} \subseteq X, \textit{ for some } i \leq m\}, \tag{1}$$

$$\overline{\sum_{i=1}^{m} P_i} X = \sim \underline{\sum_{i=1}^{m} P_i}(\sim X). \tag{2}$$

Similarly, the *boundary region* in MGRS can be extended as $Bn_{\sum_{i=1}^{m} P_i}(X) = \overline{\sum_{i=1}^{m} P_i} X \backslash \underline{\sum_{i=1}^{m} P_i} X$.

The following figure shows that the difference between Pawlak's rough set model and the multi-granulations rough set model.

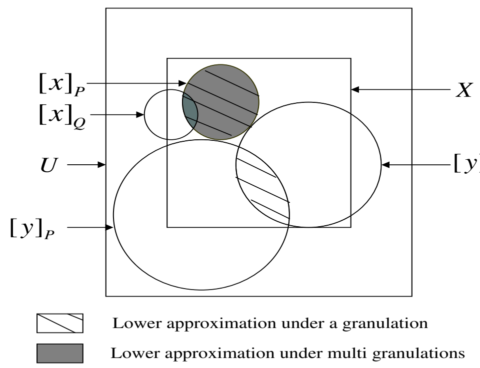

Figure 1: Difference between Pawlak's rough set model and MGRS

In Figure 1, the bias region is the lower approximation of a set X obtained by a single granulation $P \cup Q$, which are expressed by the equivalence classes in the quotient set $U/(P \cup Q)$, and the shadow region is the lower approximation of X induced by two granulations $P + Q$, which are characterized by the equivalence classes in the quotient set U/P and the quotient set U/Q together.

3 MGRS IN INCOMPLETE INFORMATION SYSTEMS

In this section, we extend the multi-granulations rough set model in complete information systems to MGRS in incomplete information systems, just incomplete MGRS.

3.1 Incomplete information systems

For an information system, any attribute domain V_a may contain special symbol "*" to indicate that the value of an attribute is unknown. Here, we assume that an object $x \in U$ possesses only one value for an attribute a, $a \in AT$. Thus, if the value of an attribute a is missing, then the real value must be from the set $V_a \backslash \{*\}$. Any domain value different from "*" will be called regular. A system in which values of all attributes for all objects from U are regular(known) is called complete, it is called incomplete otherwise [13], [14], [15], [18]. In particular, $S = (U, AT, f, D, g)$ is called an incomplete target information system if values of some attributes in AT are missing and those of all attributes in D are regular (known), where AT is called the conditional attributes and D is called the decision attributes.

Example 1: Here we employ an example to illustrate some concepts of an incomplete target information system and computations involved in our proposed incomplete MGRS. Table 1 depicts an incomplete target information system containing some information about emporium investment project. Locus, Investment and Population density are the conditional attributes of the system and Decision is the decision attribute. (In the sequel, L, I, P and D will stand for Locus, Investment, Population density and Decision, respectively.) The attribute domains are as follows: $V_L = \{good, common, bad\}$, $V_I = \{high, low\}$, $V_P = \{0.88, 0.33, 0.40, 0.37, 0.60, 0.65, 0.62\}$ and $V_D = \{Yes, No\}$.

TABLE 1
An incomplete target information system about emporium investment project

Project	Locus	Investment	Population density	Decision
e_1	common	high	0.88	Yes
e_2	bad	high	*	Yes
e_3	bad	*	0.33	No
e_4	bad	low	0.40	No
e_5	bad	low	0.37	No
e_6	bad	*	0.60	Yes
e_7	common	high	0.65	No
e_8	good	*	0.62	Yes

Let $S = (U, AT, f)$ be an incomplete information system. Each subset of attributes $P \subseteq AT$ determines a binary relation $SIM(P)$ on the universe U [13]:

$$SIM(P) = \{(u, v) \in U \times U \mid \forall a \in P, a(u) = a(v) \\ or\ a(u) = * \ or\ a(v) = *\}.$$

The relation $SIM(P)$, $P \subseteq AT$, is a tolerance relation. If the attributes $P \subseteq AT$ are numerical attributes, we define another tolerance relation as follows

$$SIM(P) = \{(u, v) \in U \times U \mid \forall a \in P, |a(u) - a(v)| \leq \delta_a \\ or\ a(u) = * \ or\ a(v) = *,\ \delta_a \geq 0\}.$$

It can be shown that $SIM(P) = \bigcup_{a \in P} SIM(\{a\})$. Let $S_P(u)$ denote the set $\{v \in U \mid (u, v) \in SIM(P)\}$. Clearly, $S_P(u)$ is the maximal set of objects which are possibly indistinguishable by P with u. Let $U/SIM(P)$ denote the family of sets $\{S_P(u) \mid u \in U\}$, called the classification or knowledge induced by the attributes P. A member $S_P(u)$ from $U/SIM(P)$ is called a tolerance class or an information granule. Note that the tolerance classes in $U/SIM(P)$ can not constitute a partition of the universe U in general. They constitute a cover of U, i.e., $S_P(u) \neq \emptyset$ for every object $u \in U$ and $\bigcup_{u \in U} S_P(u) = U$. Of particular, the *identity partition* is the cover that each of the tolerance classes contains only a singleton set and the *universal partition* is the cover that each of tolerance classes is equal to the universe set. The former is the finest cover on any nonempty set, the latter is the roughest cover on the universe U. For an incomplete target information system $S = (U, AT, f, D, g)$, if $SIM(AT) \subseteq \theta_D$, we say S is consistent, otherwise S is inconsistent [18].

We then define a partial order on the set of all classifications of U. Let $S = (U, AT, f)$ be an incomplete information system and P, $Q \in AT$. one says that P is finer than an attribute set Q (or Q is coarser than P) if and only if, $S_P(u_i) \subseteq S_Q(u_i)$ for any $i \in \{1, 2, \cdots, |U|\}$, denoted by $P \preceq Q$. If $P \preceq Q$ and $U/SIM(P) \neq U/SIM(Q)$, one says that P is strictly finer than an attribute set Q (or Q is strictly coarser than P), denoted by $P \prec Q$ [33]. In fact, $P \prec Q \Leftrightarrow S_P(u_i) \subseteq S_Q(u_i)$, $\forall i \in \{1, 2, \cdots, |U|\}$, and there exists at least one $j \in \{1, 2, \cdots, |U|\}$ such that $S_P(u_i) \subset S_Q(u_i)$.

3.2 Incomplete MGRS on two granulation spaces

Simply, we first investigate the approximation of a target set under two tolerance relations on the universe, i.e., how to approximate a target concept through using two granulation spaces.

Definition 1: Let $S = (U, AT, f)$ be an incomplete information system, $P, Q \subseteq AT$ two attribute subsets and $X \subseteq U$. A lower approximation and upper approximation of X in U are defined by the following

$$\underline{P + Q}X = \bigcup\{x \mid S_P(x) \subseteq X \ or\ S_Q(x) \subseteq X\}, \quad (3)$$

$$\overline{P+Q}X =\sim \underline{P+Q}(\sim X). \quad (4)$$

The order pair $\langle \underline{P+Q}X,\ \overline{P+Q}X\rangle$ is called a rough set of X with respect to $P+Q$. The area of uncertainty or boundary region of this rough set is defined as

$$Bn_{P+Q}(X) = \overline{P+Q}X \backslash \underline{P+Q}X.$$

One can understand the rough set approximation based on multi tolerance relations and show the difference between the incomplete MGRS and the classical rough set framework based on tolerance relation proposed by Kryszkiewicz [13] through Figure 2, Figure 3 and Example 2.

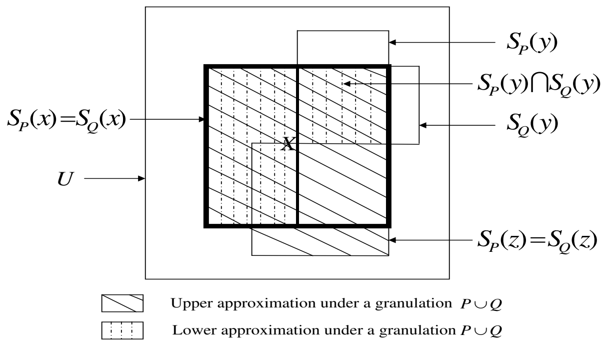

Figure 2: Set approximation in Kryszkiewicz's rough set model

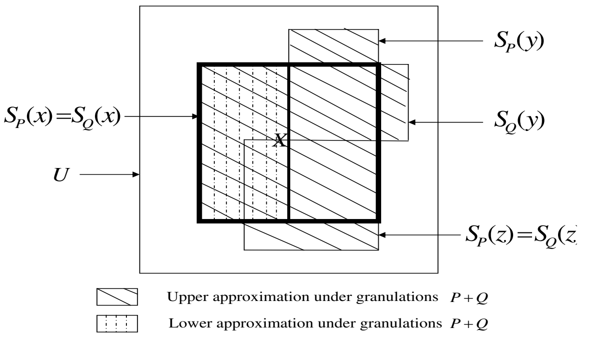

Figure 3: Set approximation in incomplete MGRS

In Figure 2, the dashed region is the lower approximation of a set X obtained by a single granulation $P\cup Q$ and the bias region is the upper approximation of X induced by the granulation $P\cup Q$ in Kryszkiewicz's incomplete rough set model. However, in Figure 3, the dashed region is the lower approximation of a set X obtained by two granulations $P+Q$ and the bias region is the upper approximation of X induced by the granulations $P+Q$ in incomplete MGRS.

Example 2: (Continued from Example 1.) Let $X = \{e_1, e_2, e_6, e_8\}$ and $\delta_P = 0.1$. Three covers can be induced from table 1 as follows

$U/SIM(L) = \{\{e_1, e_7\}, \{e_2, e_3, e_4, e_5, e_6\}, \{e_2, e_3, e_4, e_5, e_6\}, \{e_2, e_3, e_4, e_5, e_6\}, \{e_2, e_3, e_4, e_5, e_6\}, \{e_2, e_3, e_4, e_5, e_6\}, \{e_1, e_7\}, \{e_8\}\}$,

$U/SIM(P) = \{\{e_1, e_2\}, \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8\}, \{e_2, e_3, e_4, e_5\}, \{e_2, e_3, e_4, e_5\}, \{e_2, e_3, e_4, e_5\}, \{e_2, e_6, e_7, e_8\}, \{e_2, e_6, e_7, e_8\}, \{e_2, e_6, e_7, e_8\}\}$ and

$U/SIM(L\cup P) = \{\{e_1\}, \{e_2, e_3, e_4, e_5, e_6\}, \{e_2, e_3, e_4, e_5\}, \{e_2, e_3, e_4, e_5\}, \{e_2, e_3, e_4, e_5\}, \{e_2, e_6\}, \{e_7\}, \{e_8\}\}$.

From Definition 1, one can obtain that

$$\begin{aligned}\underline{L+P}X &= \bigcup\{x \mid S_L(x) \subseteq X \ or\ S_P(x) \subseteq X\}\\ &= \{e_8\} \cup \{e_1\} = \{e_1, e_8\},\\ \overline{L+P}X &=\sim \underline{L+P}(\sim X)\\ &=\sim \{\emptyset \cup \emptyset\} = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8\}.\end{aligned}$$

However, the lower approximation and the upper approximation of X in the classical rough set model based on a single tolerance relation are calculated as follows

$$\begin{aligned}\underline{L\cup P}X &= \bigcup\{x \mid S_{L\cup P}(x) \subseteq X\} = \{e_1, e_6, e_8\},\\ \overline{L\cup P}X &= \bigcup\{x \mid S_{L\cup P}(x) \cap X \neq \emptyset\}\\ &= \{e_1, e_2, e_3, e_4, e_5, e_6, e_8\}.\end{aligned}$$

Clearly, it follows from the above computations that

$$\begin{aligned}\underline{L+P}X &= \{e_1, e_8\} \subseteq \{e_1, e_6, e_8\} = \underline{L\cup P}X,\\ \overline{L+P}X &= \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8\}\\ &\supseteq \{e_1, e_2, e_3, e_4, e_5, e_6, e_8\} = \overline{L\cup P}X.\end{aligned}$$

The difference between two kinds of set approximations can be easily understood by the following theorem.

Theorem 1: Let $S = (U, AT, f)$ be an incomplete information system, $X \subseteq U$ and $P, Q \subseteq AT$ two attribute subsets. Then

1) $\underline{P+Q}X \subseteq \underline{P\cup Q}X$ and
2) $\overline{P+Q}X \supseteq \overline{P\cup Q}X$.

Proof: 1) For any $x \in \underline{P+Q}X$, from Definition 1, it follows that $S_P(x) \subseteq X$ or $S_Q(x) \subseteq X$. Hence, one has $x \in S_P(x) \cap S_Q(x)$. But $S_P(x) \cap S_Q(x) \in S_{P\cup Q}(x)$, $\forall x \in U$, therefore $x \in S_{P\cup Q}(x)$, i.e., $\underline{P+Q}X \subseteq \underline{P\cup Q}X$.

2) From the classical rough set model based on tolerance relation, we know $\overline{P\cup Q}X =\sim \underline{P\cup Q}(\sim X)$. According to the result of 1), one can obtain that $\overline{P\cup Q}X =\sim \underline{P\cup Q}(\sim X) \subseteq\sim \underline{P+Q}(\sim X) = \overline{P+Q}X$.

Corollary 1: $Bn_P(X) \subseteq Bn_{P+Q}(X)$ and $Bn_Q(X) \subseteq Bn_{P+Q}(X)$.

From the definition of set approximations under two granulation spaces, one can get the following properties of the lower approximation and the upper approximation.

Theorem 2: Let $S = (U, AT, f)$ be an incomplete information system, $X \subseteq U$ and $P, Q \subseteq AT$ two attribute subsets. Then, we have

1) $\underline{P+Q}X \subseteq X \subseteq \overline{P+Q}X$;
2) $\underline{P+Q}\emptyset = \overline{P+Q}\emptyset = \emptyset$, $\underline{P+Q}U = \overline{P+Q}U = U$;
3) $\underline{P+Q}(\sim X) =\sim \overline{P+Q}X$,
$\overline{P+Q}(\sim X) =\sim \underline{P+Q}X$;

4) $\underline{P+Q}X = \underline{P}X \cup \underline{Q}X$;
5) $\overline{P+Q}X = \overline{P}X \cap \overline{Q}X$;
6) $\underline{P+Q}X = \underline{Q+P}X$, $\overline{P+Q}X = \overline{Q+P}X$.

To establish the relationship between the approximation of a single set and that of two sets approximated through using two granulations, the following properties are given.

Theorem 3: Let $S = (U, AT, f)$ be an incomplete information system, $X, Y \subseteq U$ and $P, Q \subseteq AT$ two attribute subsets. Then, the following properties hold

1) $\underline{P+Q}(X \cap Y) = (\underline{P}X \cap \underline{P}Y) \cup (\underline{Q}X \cap \underline{Q}Y)$;
2) $\overline{P+Q}(X \cup Y) = (\overline{P}X \cup \overline{P}Y) \cap (\overline{Q}X \cup \overline{Q}Y)$;
3) $\underline{P+Q}(X \cap Y) \subseteq \underline{P+Q}X \cap \underline{P+Q}Y$;
4) $\overline{P+Q}(X \cup Y) \supseteq \overline{P+Q}X \cup \overline{P+Q}Y$;
5) $X \subseteq Y \Rightarrow \underline{P+Q}X \subseteq \underline{P+Q}Y$;
6) $X \subseteq Y \Rightarrow \overline{P+Q}X \subseteq \overline{P+Q}Y$;
7) $\underline{P+Q}(X \cup Y) \supseteq \underline{P+Q}X \cup \underline{P+Q}Y$;
8) $\overline{P+Q}(X \cap Y) \subseteq \overline{P+Q}X \cap \overline{P+Q}Y$.

3.3 Incomplete MGRS on multi granulation spaces

Based on the above conclusions, we then extend to the rough set model based on a single tolerance relation to a rough set model based on multi-granulations in the context of incomplete information systems.

Definition 2: Let $S = (U, AT, f)$ be an incomplete information system, $X \subseteq U$ and $P_1, P_2, \cdots, P_m \subseteq AT$ be m attribute subsets. We define a lower approximation of X and an upper approximation of X with respect to $P_1, P_2, \cdots, P_m$ by the following

$$\underline{\sum_{i=1}^{m} P_i} X = \bigcup \{x \mid S_{P_i}(x) \subseteq X, \textit{for some } i \leq m\}, \quad (5)$$

$$\overline{\sum_{i=1}^{m} P_i} X = \sim \underline{\sum_{i=1}^{m} P_i} (\sim X). \quad (6)$$

Similarly, the area of uncertainty or boundary region in incomplete MGRS can be represented as

$$Bn_{\underline{\sum_{i=1}^{m} P_i}}(X) = \overline{\sum_{i=1}^{m} P_i} X \backslash \underline{\sum_{i=1}^{m} P_i} X.$$

From this definition we obtain the following interpretation:

• The lower approximation of a set X with respect to $\sum_{i=1}^{m} P_i$ is the set of all objects, which can be for certain classified as X using $\sum_{i=1}^{m} P_i$ (are certainly X in view of $\sum_{i=1}^{m} P_i$).

• The upper approximation of a set X with respect to $\sum_{i=1}^{m} P_i$ is the set of all objects, which can be for possibly classified as X using $\sum_{i=1}^{m} P_i$ (are possibly X in view of $\sum_{i=1}^{m} P_i$).

• The boundary region of a set X with respect to $\sum_{i=1}^{m} P_i$ is the set of all objects, which can be classified neither as X nor as not $\sim X$ using $\sum_{i=1}^{m} P_i$.

To apply this approach in practical issues, we here present an algorithm for computing a lower approximation of a set X in this rough set model based on multi tolerance relations.

Algorithm I: Let $S = (U, AT, f)$ be an incomplete information system, $X \subseteq U$ and $P \subseteq AT$, where $P = \{P_1, P_2, \cdots, P_m\}$.

This algorithm gives the lower approximation of X by P: $\underline{\sum_{i=1}^{m} P_i} X = \bigcup \{x \mid S_{P_i}(x) \subseteq X, i \leq m\}$.

We use the following pointers:

$i = 1, 2, \cdots, m$ points to P_i,

$j = 1, 2, \cdots, |U|$ points to $S_{P_i}(u_j) \in U/SIM(P_i)$,

L records the computation of the lower approximation.

For every i and every j, we check whether or not $S_{P_i}(u_j) \cap X = S_{P_i}(u_j)$. If $S_{P_i}(u_j) \cap X = S_{P_i}(u_j)$ then we put u_j into the lower approximation of X: $L \leftarrow L \cup \{u_j\}$.

(I1) Compute m covers: $U/SIM(P_1)$, $U/SIM(P_2), \cdots$, $U/SIM(P_m)$;
(I2) Set $i \leftarrow 1$, $j \leftarrow 1$, $L = \emptyset$;
(I3) For $i = 1$ to m Do
    For $j = 1$ to $|U|$ Do
        If $S_{P_i}(u_j) \cap X = S_{P_i}(u_j)$, then
            let $L \leftarrow L \cup \{u_j\}$,
        Endif
    Endfor
    Set $j \leftarrow 1$,
  Endfor
(I4) The computation of the lower approximation X by P is completed. Output the set L.

We know that the time complexity of computing m covers is $O(m|U|^2)$. The time complexity of (I3) is also $O(m|U|^2)$ as there are $\sum_{i=1}^{m} |P_i|$ intersections $Y_i^j \cap X$ ($\leq |U| \times |U|$) to be calculated. Hence, the time complexity of Algorithm I is O($m|U|^2$).

This algorithm can be run in parallel mode to compute concurrently all corresponding covers and intersections from many attributes. Its time complexity will be O($|U|^2$). Like this idea, the algorithm for computing the upper approximation of a set also can be designed correspondingly.

Directly from Definition 2 one can obtain the following properties of the lower approximation and the upper approximation in incomplete MGRS.

Theorem 4: Let $S = (U, AT, f)$ be an incomplete information system, $X \subseteq U$ and $P_1, P_2, \cdots, P_m \subseteq AT$ be m attribute subsets. Then, the following properties hold

1) $\underline{\sum_{i=1}^{m} P_i} X = \bigcup_{i=1}^{m} \underline{P_i} X$;
2) $\overline{\sum_{i=1}^{m} P_i} X = \bigcap_{i=1}^{m} \overline{P_i} X$;
3) $\overline{\sum_{i=1}^{m} P_i} (\sim X) = \sim \underline{\sum_{i=1}^{m} P_i} X$;
4) $\underline{\sum_{i=1}^{m} P_i} (\sim X) = \sim \overline{\sum_{i=1}^{m} P_i} X$.

Proof: 1) From the formula 4) in Theorem 2, it can be easily proved.

2) From the formula (3) and 1) in this theorem, one has

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

3) From the formula (3), it is straightforward.

4) Let $X = \sim X$ in the formula (3), we have $\underline{\sum_{i=1}^m P_i}(\sim X) = \sim \overline{\sum_{i=1}^m P_i}X$.

Theorem 5: Let $S = (U, AT, f)$ be an incomplete information system, $X_1, X_2, \cdots, X_n \subseteq U$ be n subsets on U and $P_1, P_2, \cdots, P_m \subseteq AT$ be m attribute subsets. Then, the following properties hold

1) $\underline{\underline{\sum_{i=1}^m P_i}}(\bigcap_{j=1}^n X_j) = \bigcup_{i=1}^m(\bigcap_{j=1}^n \underline{P_i}X_j)$;
2) $\overline{\overline{\sum_{i=1}^m P_i}}(\bigcup_{j=1}^n X_j) = \bigcap_{i=1}^m(\bigcup_{j=1}^n \overline{P_i}X_j)$;
3) $\underline{\underline{\sum_{i=1}^m P_i}}(\bigcap_{j=1}^n X_j) \subseteq \bigcap_{j=1}^n(\underline{\underline{\sum_{i=1}^m P_i}}X_j)$;
4) $\overline{\overline{\sum_{i=1}^m P_i}}(\bigcup_{j=1}^n X_j) \supseteq \bigcup_{j=1}^n(\overline{\overline{\sum_{i=1}^m P_i}}X_j)$;
5) $\underline{\underline{\sum_{i=1}^m P_i}}(\bigcup_{j=1}^n X_j) \supseteq \bigcup_{j=1}^n(\underline{\underline{\sum_{i=1}^m P_i}}X_j)$;
6) $\overline{\overline{\sum_{i=1}^m P_i}}(\bigcap_{j=1}^n X_j) \subseteq \bigcap_{j=1}^n(\overline{\overline{\sum_{i=1}^m P_i}}X_j)$.

Theorem 6: Let $S = (U, AT, f)$ be an incomplete information system, $X_1, X_2, \cdots, X_n \subseteq U$ with $X_1 \subseteq X_2 \subseteq \cdots \subseteq X_n$ be n subsets on U and $P_1, P_2, \cdots, P_m \subseteq AT$ be m attribute subsets. Then, the following properties hold

1) $\underline{\underline{\sum_{i=1}^m P_i}}X_1 \subseteq \underline{\underline{\sum_{i=1}^m P_i}}X_2 \subseteq \cdots \subseteq \underline{\underline{\sum_{i=1}^m P_i}}X_n$;
2) $\overline{\overline{\sum_{i=1}^m P_i}}X_1 \subseteq \overline{\overline{\sum_{i=1}^m P_i}}X_2 \subseteq \cdots \subseteq \overline{\overline{\sum_{i=1}^m P_i}}X_n$.

Proof: Suppose $1 \leq i \leq j \leq n$, then $X_i \subseteq X_j$.

1) Clearly, $X_i \cap X_j = X_i$. Hence, it follows from 3) in Theorem 5 that

$$\underline{\sum_{i=1}^m P_i}X_i = \underline{\sum_{i=1}^m P_i}(X_i \cap X_j) \subseteq \underline{\sum_{i=1}^m P_i}X_i \cap \underline{\sum_{i=1}^m P_i}X_j.$$

Thus $\underline{\sum_{i=1}^m P_i}X_i = \underline{\sum_{i=1}^m P_i}X_i \cap \underline{\sum_{i=1}^m P_i}X_j$. So we have that $\underline{\sum_{i=1}^m P_i}X_i \subseteq \underline{\sum_{i=1}^m P_i}X_j$. Therefore, it follows that $\underline{\sum_{i=1}^m P_i}X_1 \subseteq \underline{\sum_{i=1}^m P_i}X_2 \subseteq \cdots \subseteq \underline{\sum_{i=1}^m P_i}X_n$.

2) Clearly, $X_i \cup X_j = X_j$. Hence, it follows from 4) in Theorem 5 that

$$\overline{\sum_{i=1}^m P_i}X_j = \overline{\sum_{i=1}^m P_i}(X_i \cup X_j) \supseteq \overline{\sum_{i=1}^m P_i}X_i \cup \overline{\sum_{i=1}^m P_i}X_j.$$

Thus $\overline{\sum_{i=1}^m P_i}X_j = \overline{\sum_{i=1}^m P_i}X_i \cup \overline{\sum_{i=1}^m P_i}X_j$. So we have that $\overline{\sum_{i=1}^m P_i}X_i \subseteq \overline{\sum_{i=1}^m P_i}X_j$. Therefore, it follows that $\overline{\sum_{i=1}^m P_i}X_1 \subseteq \overline{\sum_{i=1}^m P_i}X_2 \subseteq \cdots \subseteq \overline{\sum_{i=1}^m P_i}X_n$.

Theorem 7: Let $S = (U, AT, f)$ be an incomplete information system, $X \subseteq U$ and $P = \{P_1, P_2, \cdots, P_m\}$ with $P_1 \preceq P_2 \preceq \cdots \preceq P_m$, $\forall P_i \subseteq A$ $(i \leq m)$. Then

1) $\underline{\underline{\sum_{i=1}^m P_i}}X = \underline{P_1}X$;
2) $\overline{\overline{\sum_{i=1}^m P_i}}X = \overline{P_1}X$.

Proof: Suppose $1 \leq j \leq k \leq m$ with $P_j \preceq P_k$. From the definition of $\preceq$, we know that for any $S_{P_j}(x) \in U/SIM(P_j)$, there exits $S_{P_k}(x) \in U/SIM(P_k)$ such that $S_{P_j}(x) \subseteq S_{P_k}(x)$. Therefore, one can obtain that

$$\underline{P_j}X = \{x \mid S_{P_j}(x) \subseteq X\} \supseteq \underline{P_k}X = \{x \mid S_{P_k}(x) \subseteq X\},$$

i.e., $\underline{P_j + P_k}X = \underline{P_j}X \cup \underline{P_k}X = \underline{P_j}X$. Since $P_1 \preceq P_2 \preceq \cdots \preceq P_m$, one can get that $\underline{\sum_{i=1}^m P_i}X = \bigcup_{i=1}^m \underline{P_i}X = \underline{P_1}X$.

Similarly, we also have that

$$\overline{P_j}X = \{x \mid S_{P_j}(x) \cap X \neq \emptyset\} \supseteq \overline{P_k}X = \{x \mid S_{P_k}(x) \cap X \neq \emptyset\},$$

i.e., $\overline{P_j + P_k}X = \overline{P_j}X \cap \overline{P_k}X = \overline{P_j}X$. Hence, one can obtain that $\overline{\sum_{i=1}^m P_i}X = \bigcap_{i=1}^m \overline{P_i}X = \overline{P_1}X$.

3.4 Several elementary measures in incomplete MGRS

In this section, we investigate several elementary measures in incomplete MGRS and their properties.

Uncertainty of a set (category) is due to the existence of a borderline region. The bigger the borderline region of a set is, the lower the accuracy of the set is (and vice versa). To more precisely express this idea, we introduce the accuracy measure to incomplete MGRS as follows.

Definition 3: Let $S = (U, AT, f)$ be an incomplete information system, $X \subseteq U$ and $P = \{P_1, P_2, \cdots, P_m\}$, $\forall P_i \subseteq AT$. Approximation measure of X by P is defined as

$$\alpha_P(X) = \frac{|\underline{\underline{\sum_{i=1}^m P_i}}X|}{|\overline{\overline{\sum_{i=1}^m P_i}}X|}, \tag{7}$$

where $X \neq \emptyset$, $|X|$ denotes the cardinality of set X.

From this definition one can derive the following theorem.

Theorem 8: Let $S = (U, AT, f)$ be an incomplete information system, $X \subseteq U$, $P = \{P_1, P_2, \cdots, P_m\}$, $\forall P_i \subseteq AT$, and $P' \subseteq P$ a subset of P. Then

$$\alpha_P(X) \geq \alpha_{P'}(X) \geq \alpha_{P_i}(X) \ (i \leq m).$$

Proof: Since $P' \subseteq P$, it follows from Definition 2 that

$$\bigcup_{i=1}^m \underline{\underline{P_i}}X \supseteq \bigcup_{P_i \in P,\ P_i \notin P'} \underline{\underline{P_i}}X \text{ and } \bigcap_{i=1}^m \overline{\overline{P_i}}X \subseteq \bigcap_{P_i \in P,\ P_i \notin P'} \overline{\overline{P_i}}X.$$

Then, it is clear that $|\bigcup_{i=1}^m \underline{P_i}X| \geq |\bigcup_{P_i \in P,\ P_i \notin P'} \underline{P_i}X|$ and $|\bigcap_{i=1}^m \overline{P_i}X| \leq |\bigcap_{P_i \in P,\ P_i \notin P'} \overline{P_i}X|$. Hence,

$$\begin{aligned} \alpha_P(X) &= \frac{|\underline{\underline{\sum_{i=1}^m P_i}}X|}{|\overline{\overline{\sum_{i=1}^m P_i}}X|} = \frac{|\bigcup_{i=1}^m \underline{P_i}X|}{|\bigcap_{i=1}^m \overline{P_i}X|} \\ &\geq \frac{|\bigcup_{P_i \in P,\ P_i \notin P'} \underline{P_i}X|}{|\bigcap_{P_i \in P,\ P_i \notin P'} \overline{P_i}X|} = \frac{|\underline{\underline{\sum_{P_i \in P,\ P_i \notin P'}}} P_iX|}{|\overline{\overline{\sum_{P_i \in P,\ P_i \notin P'}}} P_iX|} \\ &= \alpha_{P'}(X). \end{aligned}$$

Similarly, we have $\alpha_{P'}(X) \geq \alpha_{P_i}(X)$ $(i \leq m)$. Thus, we have $\alpha_P(X) \geq \alpha_{P'}(X) \geq \alpha_{P_i}(X)$ $(i \leq m)$.

Theorem 8 shows that the approximation measure of a target concept enlarges as the number of granulations for describing the concept becomes much bigger.

Example 3: (Continued from Example 2.) Suppose $A = \{L,\ P\}$. Computing the approximation measures, it follows that

$\alpha_A(X) = \frac{|\underline{L+P}X|}{|\overline{L+P}X|} = \frac{1}{4}$,
$\alpha_L(X) = \frac{|\underline{L}X|}{|\overline{L}X|} = \frac{1}{8}$ and
$\alpha_P(X) = \frac{|\underline{P}X|}{|\overline{P}X|} = \frac{1}{8}$.

Clearly, it follows from the above computations that $\alpha_A(X) > \alpha_L(X)$ and $\alpha_A(X) > \alpha_P(X)$.

In particular, we have $\alpha_A(X) > \alpha_L(X)$ if $L \preceq P$. It can be easily derived by Theorem 7.

Note that the approximation measure of a target concept approximated by using multi granulations is always much better than that approximated by using a single granulation, which is suitable for more precisely characterizing a target concept and problem solving according to user requirements.

Definition 4: Let $S = (U, AT, f, D, g)$ an incomplete target information system, $U/IND(D)$ be a decision induced by the decision attributes D and $P = \{P_1, P_2, \cdots, P_m\}$ be m attribute sets. Approximation quality of D by P, also called a degree of dependency, is defined as

$$\gamma(P, D) = \frac{\sum\{|\underline{\sum_{i=1}^m P_i}Y| : Y \in U/IND(D)\}}{|U|}. \quad (8)$$

This measure can be used to evaluate the deterministic part of the rough set description of $U/IND(D)$ by counting those objects which can be re-classified to blocks of $U/IND(D)$ with the knowledge given by $\sum_{i=1}^m P_i$. As a result of the above definition, we come to the following two theorems.

Theorem 9: Let $P = \{P_1, P_2, \cdots, P_m\}$ be m attribute sets and D_1, D_2 with $D_1 \preceq D_2$ two decisions, then $\gamma(\sum_{i=1}^m P_i,\ D_1) \leq \gamma(\sum_{i=1}^m P_i,\ D_2)$.

Theorem 10: Let $P = \{P_1, P_2, \cdots, P_m\}$ be m attribute sets and D a decision. If $P' \subseteq P$, then $\gamma(\sum_{i=1}^m P_i,\ D) \geq \gamma(\sum_{P_i \in P'} P_i,\ D) \geq \gamma(P_i,\ D)$.

Gediga and Düntsch [6] introduced a simple statistic $\pi(R, X) = \frac{|\underline{R}X|}{|X|}$ for the precision of (deterministic) approximation of $X \subseteq U$ given $U/IND(R)$, which is not affected by the approximation of $\sim X$. This is just the relative number of elements in X which can be approximated by R, clearly, $\pi(R, X) \geq \alpha(R, X)$. It is important to point out, that $\pi(R, X)$ requires complete knowledge of X, whereas α does not, since the latter uses only the rough set $(\underline{R}X, \overline{R}X)$. In incomplete MGRS, it can be extended to be the formula

$$\pi(\sum_{i=1}^m P_i, X) = \frac{|\underline{\sum_{i=1}^m P_i}X|}{|X|}. \quad (9)$$

It is clear that $\pi(\sum_{i=1}^m P_i, X) \geq \alpha(\sum_{i=1}^m P_i, X)$. In fact, this measure denotes the relative number of objects in X which can be approximated by $\sum_{i=1}^m P_i$. Easily, the following theorem can be proved.

Theorem 11: Let $P = \{P_1, P_2, \cdots, P_m\}$ be m attribute sets and X a target concept. If $P' \subseteq P$, then $\pi(\sum_{i=1}^m P_i,\ X) \geq \pi(\sum_{P_i \in P'} P_i,\ X) \geq \pi(P_i,\ X)$.

3.5 Experimental analysis

In the following, through experimental analyses, we illustrate the deference between the incomplete MGRS and Kryszkiewicz's rough set model. We have downloaded four public data sets (incomplete target information systems) with practical applications from UCI Repository of machine learning databases, which are described in Table 2.

Here, we compare the degree of dependency in incomplete MGRS with that in Kryszkiewicz's rough set model on these two practical data sets. The comparisons of values of two measures with the numbers of features in these two data sets are shown in Figures 4-7.

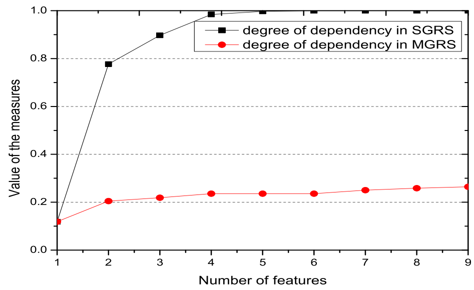

Figure 4: Variation of the two degrees of dependency with the numbers of features (data set breast-cancer-wisconsin)

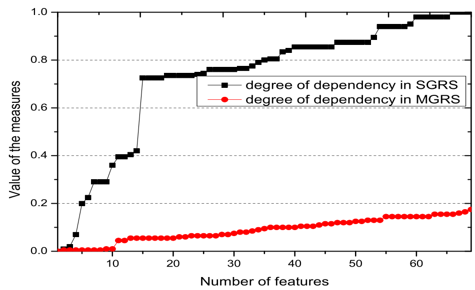

Figure 5: Variation of the two degrees of dependency with the numbers of features (data set audiology)

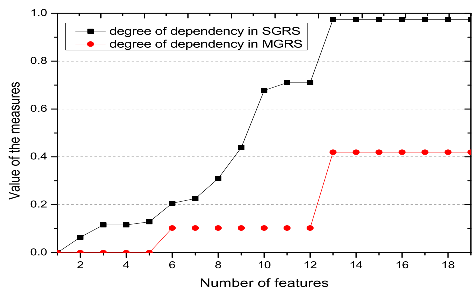

Figure 6: Variation of the two degrees of dependency with the numbers of features (data set hepatitis)

TABLE 2
Data sets description

Data sets	Samples	Numerical features	Symbolic features	Decision classes
breast-cancer-wisconsin	699	0	9	2
audiology	200	0	69	24
hepatitis	155	6	13	2
crx	690	6	9	2

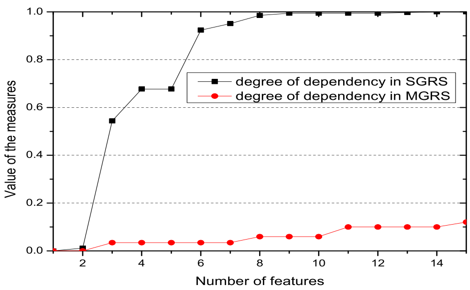

Figure 7: Variation of the two degrees of dependency with the numbers of features (data set crx)

In Figures 4-7, the term MGRS is the incomplete multi-granulations rough set framework proposed in this paper, and the term SGRS is Kryszkiewicz's rough set model. It can be seen from Figures 4-7 that the value of the degree of dependency in incomplete MGRS is not bigger than that in Kryszkiewicz's rough set model for the same number of selected features, and this value increases as the number of selected features becomes bigger in the same data set. In particular, from Figure 6, it is easy to see that the values of the degree of dependency in incomplete MGRS are equal to zero when the number of features falls in between 1 and 5. In this situation, the lower approximation of the target decision equals an empty set in the incomplete decision table. In essence, it is because that the tolerance classes induced by a singleton attribute are all coarser than those induced by all attributes. One can draw the same conclusion from other figures. Therefore, although the incomplete MGRS can not obtain more bigger the approximation measure and the degree of dependency than Kryszkiewicz's rough set model, this approach can be used to concept representation, rule extraction and data analysis from incomplete data sets under multi granulations on the basis of keeping original granulation structure. Further illustration and applications are shown in Section V.

4 Attribute Reduction in incomplete MGRS

Reduct is a minimal attribute subset of the original data which is independent and has the same discernibility power as all of the attributes in classical rough set framework. Obviously reduction is a feature subset selection process, where the selected feature subset not only retains the representation power, but also has minimal redundancy [22], [23]. In this section, we deal with attribute reduction in incomplete MGRS.

We first introduce the notions of two approximation distribution functions. Let $S = (U, AT, f, D, g)$ be an incomplete target information system, $P \subseteq AT$, and the decision $U/IND(D) = \{Y_1, Y_2, \cdots, Y_r\}$. Low approximation distribution function and upper approximation function are defined as

$$\underline{D}_P = (\underline{\sum_{P_i\in P} P_i}Y_1, \underline{\sum_{P_i\in P} P_i}Y_2, \cdots, \underline{\sum_{P_i\in P} P_i}Y_r),$$
$$\overline{D}_P = (\overline{\sum_{P_i\in P} P_i}Y_1, \overline{\sum_{P_i\in P} P_i}Y_2, \cdots, \overline{\sum_{P_i\in P} P_i}Y_r).$$

Through using these two approximation distribution functions, three new reducts can be defined in the following, which are lower approximation reduct, upper approximation reduct and approximation reduct.

Definition 5: Let $S = (U, AT, f, D, g)$ be an incomplete target information system and P is a non-empty subset of AT.

1) If $\underline{D}_P = \underline{D}_{AT}$, we say that P is a *lower approximation consistent set* of S. If P is a lower approximation consistent set and no proper subset of P is lower approximation consistent, then P is called a *lower approximation reduct* of S.

2) If $\overline{D}_P = \overline{D}_{AT}$, we say that P is a *upper approximation consistent set* of S. If P is a upper approximation consistent set and no proper subset of P is upper approximation consistent, then P is called a *upper approximation reduct* of S.

3) If P is not only a lower approximation reduct but also a upper approximation reduct, then P is called an *approximation reduct* of S.

It is easy to prove that a upper approximation consistent set must be a lower approximation consistent set. However the converse relationship cannot be satisfied in an inconsistent incomplete target information system. From the above definition, it is clear that P is a lower approximation consistent set if and only if P is a upper approximation consistent set in a consistent incomplete target information system. In particular, if $U/IND(D) = X$, we can regard P as a lower approximation reduct, a upper approximation reduct and an approximation reduct of a target concept X, respectively.

Let **A** be the set of all lower approximation reducts, **B** be the set of all upper approximation reducts, it is obvious that the approximation reducts $\mathbf{C} = \mathbf{A} \cap \mathbf{B}$.

Suppose that $S = (U, AT, f, D, g)$ be an incomplete target information system, where $U = \{e_1, e_2, \cdots,$

$e_{|U|}\}$, $AT = \{P_1, P_2, \cdots, P_{|AT|}\}$ and $U/IND(D) = \{Y_1, Y_2, \cdots, Y_r\}$. We denote all lower approximation reducts of $Y \in U/IND(D)$ by $\mathbf{A}(Y)$, all upper approximation reducts of $Y \in U/IND(D)$ by $\mathbf{B}(Y)$ and all approximation reducts of $Y \in U/IND(D)$ by $\mathbf{C}(Y)$, respectively. And we call $Core(\mathbf{A}(Y))$ the lower approximation core of Y, $Core(\mathbf{B}(Y))$ the upper approximation core of Y and $Core(\mathbf{C}(Y))$ the approximation core of Y, respectively.

Theorem 12: Let $S = (U, AT, f, D, g)$ be an incomplete target information system and $U/IND(D) = \{Y_1, Y_2, \cdots, Y_r\}$. Then

$$\mathbf{A} = \bigcap_{k=1}^{r} \mathbf{A}(Y_k) \text{ and } \mathbf{B} = \bigcap_{k=1}^{r} \mathbf{B}(Y_k).$$

We call $Core(\mathbf{A}) = \bigcap \mathbf{A}_i(\mathbf{A}_i \in \mathbf{A})$, $Core(\mathbf{B}) = \bigcap \mathbf{B}_i(\mathbf{B}_i \in \mathbf{B})$ and $Core(S) = \bigcap \mathbf{C}_i(\mathbf{C}_i \in \mathbf{C})$ the lower approximation core, the upper approximation core and the core of an incomplete target information system S, respectively.

Theorem 13: Let $S = (U, AT, f, D, g)$ be an incomplete target information system and $U/IND(D) = \{Y_1, Y_2, \cdots, Y_r\}$. Then

$$Core(\mathbf{A}) = \bigcap_{k=1}^{r} Core(\mathbf{A}(Y_k)),$$
$$Core(\mathbf{B}) = \bigcap_{k=1}^{r} Core(\mathbf{B}(Y_k)).$$

Clearly, we have that $Core(S) = Core(\mathbf{A}) \cap Core(\mathbf{B})$. In fact, the core is indispensable attribute to construct an approximation reduct. Figure 8 shows the relationship the approximation reducts and the approximation core of a target information system.

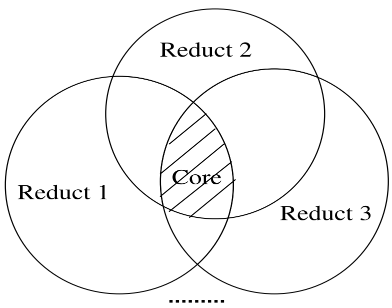

Figure 8: Relationship between approximation reducts and approximation core

Then, we discuss how to find attribute reducts from an incomplete target information system in the framework of incomplete MGRS. Let $S = (U, AT, f, d, g)$ be an incomplete target information system, where $U = \{e_1, e_2, \cdots, e_U\}$, $AT = \{P_1, P_2, \cdots, P_{AT}\}$ and $U/IND(\{d\}) = \{Y_1, Y_2, \cdots, Y_r\}$. In the framework of incomplete MGRS, we then design for computing all lower approximation reducts-that is, all subsets $AT_0 : AT_{01}, AT_{02}, \cdots, AT_{0s}$ of AT such that:

1) $\underline{d}_{AT_0} = \underline{d}_{AT}$ and
2) if $AT^{'} \subset AT_0$, then $\underline{d}_{AT_0} \neq \underline{d}_{AT}$.

Algorithm II: This algorithm gives all lower approximation reducts of the incomplete target information system S. (Similar to the idea in Algorithm K in [7].)

Let us denote the binomial coefficients by $C_{|AT|}^{k} = |AT|!/k!(|AT|-k)!$.

(1) Let us denote $C_{|AT|}^{1} = |AT|$ singletons, one-attribute subsets, by
$AT_{11} = \{P_1\}$, $AT_{12} = \{P_2\}$, $\cdots$, $AT_{1j} = \{P_j\}$, $\cdots$, $AT_{1C_{|AT|}^{1}} = \{P_{|AT|}\}$.

(2) Let us denote $C_{|AT|}^{2} = |AT|(|AT|-1)/2!$ two-attribute subsets by
$AT_{21} = \{P_1, P_2\}$, $\cdots$, $AT_{2j} = \{P_1, P_j\}$, $\cdots$, $AT_{2C_{|AT|}^{2}} = \{P_{|AT|-1}, P_{|AT|}\}$.

(3) Generally, let us denote $C_{|AT|}^{k} = |AT|!/k!(|AT|-k)!$ k-attribute subsets by
$AT_{k1} = \{P_1, P_2, \cdots, P_k\}$, $\cdots$, AT_{kj}, $\cdots$, $AT_{kC_{|AT|}^{k}} = \{P_{|AT|-k+1}, \cdots, P_{|AT|-1}, P_{|AT|}\}$.

(4) Notice that $C_{|AT|}^{|AT|} = 1$ and $|AT|$-attribute subset is $AT_{|AT|1} = AT$.

The algorithm is to search subsets of AT as follows: singletons, two-attribute subsets, $\cdots$, $t-$attribute subsets, and so on. Continue up to the unique $|AT|$-attribute subset AT itself.

We use the following variables:

- s–the number of the lower approximation reducts we have already found,
- t–counting from 1 to s,
- k–we are currently searching k-attribute subset AT_{kj},
- j–we are currently searching the jth subset AT_{kj} in all k-attribute subsets $AT_{k1}, \cdots, AT_{kj}, \cdots, AT_{kC_{|AT|}^{k}}$.

(II1) Set $j \leftarrow 1$, $s \leftarrow 0$, $k \leftarrow 1$;
(II2) While $k \leq |AT|$ Do
    $j \leftarrow 1$;
    While $j \leq C_{|AT|}^{k}$ Do
        for $t = 1$ to s Do
            If $AT_{0t} \subset AT_{kj}$, then break;
            Endif
        Endfor
        if $\underline{d}_{AT_{kj}} = \underline{d}_{AT}$, then
            $s \leftarrow s + 1$, $AT_{0s} \leftarrow AT_{kj}$;
        Endif
        $j \leftarrow j + 1$;
    Endwhile
    $k \leftarrow k + 1$;
  Endwhile
(II3) Output $AT_{01}, AT_{02}, \cdots, AT_{0s}$ (s lower approximation reducts).

The time complexity of this algorithm for all lower approximation reducts is exponential since it checks all subsets in 2^{AT}, and $|2^{AT}| = 2^{|AT|}$. We know that the time complexity of computing $|AT|$ covers is $O(|AT||U|^2)$ and the time complexity of computing a lower approximation of every $Y \in U/IND(\{d\})$ by AT_{kj} ($k \leq |AT|$) is $O(|AT||U|^3)$. Thus, the time complexity of Algorithm II is

$$2^{|AT|} \times O(|AT||U|^2 + |AT||U|^3) = O(2^{|AT|}|AT||U|^3).$$

Through Algorithm II, one can obtain that the attribute set $\{L, P\}$ is only one lower approximation reduct of Table 1. However, the time complexity of Algorithm II is exponential so that it cannot be applied efficiently in practical applications. To reduce the time complexity of computing an approximation reduct, we introduce in the following two heuristic functions, which are an important measure of lower approximation and an important measure of upper approximation.

Let $S = (U, AT, f, D, g)$ be an incomplete target information system and P a non-empty subset of AT. Given a condition attribute $a \in P$ and $Y \in U/IND(D)$, we first give two preliminary definitions in the following, which will be helpful for constructing heuristic functions.

Definition 6: We say that a is lower approximation significant in P with respect to X if

$$\underline{\sum_{i=1}^{|P|} P_i}X \supset \underline{\sum_{i=1,P_i\neq a}^{|P|} P_i}X,\ (P_i \in P)$$

and a is not lower approximation significant in P with respect to X if

$$\underline{\sum_{i=1}^{|P|} P_i}X = \underline{\sum_{i=1,P_i\neq a}^{|P|} P_i}X,\ (P_i \in P),$$

where $|P|$ is the cardinality of attribute set P.

Definition 6 shows that if a is lower approximation significant with respect to X, then the lower approximation of X will become smaller; if a is not lower approximation significant with respect to X, then the lower approximation of X will keep unchanged.

Definition 7: We say that a is upper approximation significant in P with respect to X if

$$\overline{\sum_{i=1}^{|P|} P_i}X \subset \overline{\sum_{i=1,P_i\neq a}^{|P|} P_i}X,\ (P_i \in P)$$

and a is not lower approximation significant in P with respect to X if

$$\overline{\sum_{i=1}^{|P|} P_i}X = \overline{\sum_{i=1,P_i\neq a}^{|P|} P_i}X,\ (P_i \in P),$$

where $|P|$ is the cardinality of attribute set P.

Analogous to Definition 6, Definition 7 states that if a is upper approximation significant with respect to X, then the upper approximation of X will become bigger; if a is not upper approximation significant with respect to X, then the upper approximation of X will keep unchanged.

Through these two definitions, we easily construct two heuristic functions. An important measure of lower approximation of $P \subseteq AT$ with respect to D is defined as

$$S_P(D) = \frac{\sum\limits_{Y\in U/D} |\underline{\sum\limits_{i=1}^{m} P_i}Y \backslash \underline{\sum\limits_{i=1,P_i\notin P}^{m} P_i}Y|}{|U|} \quad (10)$$

and an important measure of upper approximation of $P \subseteq AT$ with resect to D is defined as

$$S^P(D) = \frac{\sum\limits_{Y\in U/D} |\overline{\sum\limits_{i=1,P_i\notin P}^{m} P_i}Y \backslash \overline{\sum\limits_{i=1}^{m} P_i}Y|}{|U|}. \quad (11)$$

In particular, when $P = \{a\}$, $S_a(D)$ and $S^a(D)$ can be regard as the importance measure of lower approximation and the importance measure of upper approximation of the attribute $a \in AT$ with respect to D, respectively.

From Algorithm I, we know that the time complexity of computing the lower approximation of Y by $P = \{P_1, P_2, \cdots, P_m\}$ is $O(m|U|^2)$. For computing the measure of important of a, we need to calculate the lower approximations for at most $|U|$ times. Therefore, the time complexity of computing an important measure of lower/upper approximation of an attribute with respect to D is $O(m|U|^3)$.

From the above denotations, we come to the following conclusions:

- $S_P(D) \geq 0$ and $S^P(D) \geq 0$,
- P is not lower approximation significant with respect to D iff $S_P(D) = 0$ and
- P is not upper approximation significant with respect to D iff $S^P(D) = 0$.

We then provide a heuristic algorithm based on the importance measure of lower approximation of a condition attribute with respect to the decision attribute d to find a lower approximation reduct in an incomplete target information system.

Algorithm III: Let $S = (U, AT, f, d, g)$ be a complete target information system, where $U = \{e_1, e_2, \cdots, e_{|U|}\}$, $AT = \{P_1, P_2, \cdots, P_{|AT|}\}$ and $U/IND(\{d\}) = \{Y_1, Y_2, \cdots, Y_r\}$.

This algorithm find a lower approximation reduct through using a heuristic information.

We use the following variables:

- AT_0–It is used to record a lower approximation reduct,
- i–we are currently searching the ith condition attribute $AT_i^{'}$ in the sequence given.

(III1) Compute $|AT|$ covers and a decision partition $U/IND(\{d\})$;
(III2) Sort $AT = \{P_1^{'}, P_2^{'}, \cdots, P_{|AT|}^{'}\}$, where $S_{P_i^{'}}(d) \geq S_{P_{i+1}^{'}}(d)$;
(III3) Set $i \leftarrow 1$, $AT_0 = \emptyset$.
(III4) If $\underline{d}_{AT_0} \neq \underline{d}_{AT}$, then
$AT_0 \leftarrow AT_0 \cup P_i^{'}$,
$i \leftarrow i + 1$;
(III5) Found a lower approximation reduct: AT_0. Output the set AT_0.

The time complexity of this algorithm for computing $|AT|$ covers and a decision partition $U/IND(\{d\})$ is $O(|AT||U|^2)$. The time complexity of computing $|AT|$

TABLE 3
An incomplete evaluation table about venture investment

U	E_1	E_2	E_3	E_4	E_5	D
x_1	2	3	3	2	2	High
x_2	1	3	3	2	2	High
x_3	1	1	1	1	1	Low
x_4	1	1	1	1	1	Low
x_5	1	1	1	1	1	Low
x_6	$*$	2	1	2	3	High
x_7	2	2	2	2	2	Low
x_8	3	2	2	3	3	High
x_9	2	3	2	3	1	High
x_{10}	1	1	$*$	1	$*$	Low

importance measures is $O(|AT||U|^3)$ and the time complexity of sorting is $O(|AT|log_2|AT|)$. And the time complexity for running $|AT|$ comparisons $\underline{d}_{AT_0} = \underline{d}_{AT}$ is $O(|AT||U|^3)$. Thus, the time complexity of Algorithm III is

$$O(|AT||U|^2 + |AT||U|^3 + |AT|log_2|AT| + |AT||U|^3) \\ = O(|AT||U|^3).$$

Through Algorithm III a lower approximation reduct can be found, which keeps the lower approximation distribution function of this incomplete target information system. Analogously, we can design a heuristic algorithm to find a upper approximation reduct of an incomplete target information system through using a heuristic function $S^P(D)$.

5 An application for venture investment

Venture capital has become in increasingly important source of financing for new companies, particularly when such companies are operating on the frontier of emerging technologies and markers. It plays an essential role in the entrepreneurial process. For an investor or decision maker, he may need to adopt a better one from some possible investment projects or find some directions from existing successful investment projects before investing. The purpose of this section is, through a venture investment issue, to illustrate the mechanism of incomplete MGRS and its applications.

Let us consider a real investment issue of a venture investment company (Here we conceal the company's name and the details of investment projects). There are ten investment projects x_i $(i = 1, 2, \cdots, 10)$ can be considered, which are evaluated by five evaluation experts. Venture level is classified to three classes 1, 2 and 3. The bigger the value of venture level is, and the higher the venture of investment project is. Table 3 is an incomplete evaluation table about venture investment given by these five experts, in which the symbol "$*$" means that an expert can not decide the venture level of a project. In the evaluation process, each of evaluation experts makes a decision independently, i.e., one does not perform intersection operations between any two evaluation results. For this situation, the classical Kryszkiewicz's rough set model will be helpless. In the following, we apply incomplete MGRS proposed by this paper for decision-making.

From Table 3, it is easy to see that $U/IND(D) = \{\{x_1, x_2, x_6, x_8, x_9\}, \{x_3, x_4, x_5, x_7, x_{10}\}\}$. Suppose that $Y_1 = \{x_1, x_2, x_6, x_8, x_9\}$ and $Y_2 = \{x_3, x_4, x_5, x_7, x_{10}\}$.

To acquire certain decision rules, we only calculate the lower approximation of the decision D with respect to the five granulation spaces determined by the five experts. It easily follows from Definition 2 that

$$\underline{D}_{AT} = \{\{x_1, x_2, x_8, x_9\}, \{x_3, x_4, x_5, x_{10}\}\}.$$

From Definition 5, one can obtain the following lower approximation reducts of Table 3, which is as follows

$$\mathbf{A}(D) = \{\{E_1, E_2\}, \{E_2, E_4\}\}.$$

Thus, there is only a core E_2. That is to say, the decision given by the second expert is indispensable.

From the above two reducts, one can extract two groups of certain decision rules in the following.

$(E_1 = 3) \vee (E_2 = 3) \Rightarrow (D = High)$,
$(E_2 = 1) \Rightarrow (D = Low)$;
and
$(E_2 = 3) \vee (E_4 = 3) \Rightarrow (D = High)$,
$(E_2 = 1) \vee (E_4 = 1) \Rightarrow (D = Low)$.

In addition, from the formula (10) one can calculate the important measure of lower approximation of each expert's decision, which listed as follows.

$S_{E_1}(D) = 0.2$, $S_{E_2}(D) = 0.7$, $S_{E_3}(D) = 0$, $S_{E_4}(D) = 0$ and $S_{E_5}(D) = 0$.

Hence, one can know that the decision of expert E_2 is the most important, E_2 takes second place, and so on. Through using the sequence and Algorithm III, we can obtain one of reducts from Table 3, which is $\{E_2, E_1\}$.

Remark. The incomplete multi-granulations rough set model (incomplete MGRS) does not attempt to keep or reduce the uncertainty induced by the classical Kryszkiewicz's model, but aims to concept representation and rule extraction on the basis of keeping the original granulation structures. The incomplete MGRS has several useful applications. (1) It can deal with intelligent decision-making under multi granulations. For example, the above evaluation issue demands that each of evaluation experts makes a decision independently, i.e., one does not perform intersection operations between any two evaluation results. (2) To extract decision rules from distributive decision systems using rough set approaches, the incomplete MGRS can largely reduce the time complexity of rule extraction when the increase of uncertainty is tolerable, in which there is no need to perform the intersection operations in between all the sites.

6 Conclusion

The contribution of this paper is twofold. On one side, the incomplete single-granulation rough set theory has been extended and an incomplete multi-granulations rough set model has been obtained. In this extension, the approximations of sets are defined by using multi tolerance relations on the universe. These tolerance relations can be chosen according to user requirements or targets of problem solving. The theoretical analyses show that some properties of the original incomplete rough set model become special instances of incomplete MGRS. Under the incomplete MGRS, we also have developed several important measures, such as the accuracy measure, the quality of approximation and the precision of approximation. On the other side, to acquire brief representation for the approximation of a target decision, the attribute reduction has been discussed in incomplete information systems. A concept of approximation reduct has been used to characterize the smallest attribute subset that preserves the lower approximation and upper approximation of all decision classes in incomplete MGRS. Two key attribute reduction algorithms have been designed, which will be helpful for applying this theory in practical issues. The multi-granulations rough set framework proposed in this study maybe lead to a tool for more widely applying the incomplete rough set theory in real-world applications. The incomplete MGRS can be applied for concept representation, rule extraction and data analysis from incomplete data set under multi-granulations spaces, and has much wider applicability ranges than Kryszkiewicz's rough set model.

Acknowledgment

This work was supported by the national natural science foundation of China (No. 60773133, No. 70471003, No. 60573074), the high technology research and development program of China (No. 2007AA01Z165), the national key basic research and development program of China (973) (No. 2007CB311002), SRG: 7001977 of City University of Hong Kong.

References

[1] B. Apolloni, A. Brega, D. Malchiodi, G. Palmas and A.M. Zanaboni, "Learning rule representations from data," *IEEE Trans. Syst. Man. CY. A*, vol. 36, no. 5, pp. 1010-1028, 2006.
[2] A. Bargiela and W. Pedrycz, "Granular mappings," *IEEE Trans. Syst. Man. CY. A*, vol. 35, no. 2, pp. 292-297, 2005.
[3] S. Dick, A. Schenker, W. Pedrycz and A. Kandel, "Regranulation: A granular algorithm enabling communication between granular worlds," *Inform. Sci.*, vol. 177, pp. 408-435, 2007.
[4] I. Düntsch and G. Gediga, "Uncertainty measures of rough set prediction," *Artif. Intell.*, vol. 106, pp. 109-137, 1998.
[5] D. Dübois and H. Prade, "Rough fuzzy sets and fuzzy rough sets", *Int. J. Gen. Syst.*, vol. 17, pp. 191-209, 1990.
[6] G. Gediga and I. Düntsch, "Rough approximation quality revisited," *Artif. Intell.*, vol. 132, pp. 219-234, 2001.
[7] J. W. Guan and D. A. Bell, "Rough computational methods for information systems," *Artif. Intell.*, vol. 105, pp. 77-103, 1998.
[8] T.P. Hong, L. Tseng, and S. Wang, "Learning rules from incomplete training examples by rough sets," *Expert Syst. Appl.*, vol. 22, no. 4, pp. 285-293, 2002.
[9] Q. Hu, D. Yu, and Z. Xie, "Reduction algorithms for hybrid data based on fuzzy rough set approches," in *Proc. 2004 Int. Conf. Machine Learning and Cybernetics.*, pp. 1469-1474, 2004.
[10] Q. Hu, D. Yu, Z. Xie, and J. Liu, "Fuzzy probabilistic approximation spaces and their information measures," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 2, pp. 191-201, 2006.
[11] R. Jensen and Q. Shen, "Fuzzy-rough sets assisted attribute selection", *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 1, pp. 73-89, 2007.
[12] G. Jeon, D. Kim and J. Jeong, "Rough Sets Attributes Reduction Based Expert System in Interlaced Video Sequences," *IEEE Trans. Consum. Electr.*, vol. 52, no. 4, pp. 1348-1355, 2006.
[13] M. Kryszkiewicz, "Rough set approach to incomplete information systems," *Inform. Sci.*, vol. 112, pp. 39-49, 1998.
[14] M. Kryszkiewicz, "Rules in incomplete information systems," *Inform. Sci.*, vol. 113, pp. 271-292, 1999.
[15] Y. Leung and D. Y. Li, "Maximal consistent block technique for rule acquisition in incomplete information systems," *Inform. Sci.*, vol. 153, pp. 85-106, 2003.
[16] D. Li, B. Zhang, and Y. Leung, "On knowledge reduction in inconsistent decision information systems," *Int. J. Uncertainty, Fuzziness Knowledge-Based Syst.*, vol. 12, no. 5, pp. 651-672, 2004.
[17] J. Y. Liang, C. Y. Dang, K. S. Chin and C. M. Yam Richard, "A new method for measuring uncertainty and fuzziness in rough set theory," *Int. J. Gen. Syst.*, vol. 31, no. 4, pp. 331-342, 2002.
[18] J. Y. Liang and D. Y. Li, *Uncertainty and Knowledge Acquisition in Information Systems*, Science Press, Beijing, China, 2005.
[19] J. Y. Liang and Y. H. Qian, "Axiomatic approach of knowledge granulation in information systems", *Lecture Notes. Artif. Intell.*, vol. 4304, pp. 1074-1078, 2006.
[20] J. Y. Liang and Y. H. Qian, "Information granules and entropy theory in information systems," *Sci. China. Ser. F.*, vo. 51, no. 9, pp. 1-18, 2008.
[21] T. Y. Lin, "From rough sets and neighborhood systems to information granulation and computing in words," *Proc. Eur. Congr. Intelligent Techniques and Soft Computing*, pp. 1602-1606, Sep. 8-12, 1997.
[22] H. S. Nguyen, "Approximate boolean reasoning: foundations and applications in data mining", *Lecture Notes. Comput. Sci.*, vol. 4100, pp. 334-506, 2006.
[23] S. H. Nguyen and H. S. Nguyen, "Some efficient algorithms for rough set methods", *Proc. 6th Int. Conf. Information Processing and Management of Uncertainty on Knowledge Based Systems*, vol. III, pp. 1451-1456, 1996.
[24] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data, System Theory, Knowledge Engineering and Problem Solving*, Vol. 9, Kluwer, Dordrecht, 1991.
[25] Z. Pawlak, "Rough sets, decision algorithms and Bayes's theorem," *Eur. J. Oper. Res.*, vol. 136, no. 1, pp. 181-189, 2002.
[26] Z. Pawlak and A. Skowron, "Rudiments of rough sets," *Inform. Sci.*, vol. 177, pp. 3-27, 2007.
[27] Z. Pawlak and A. Skowron, "Rough sets: some extensions," *Inform. Sci.*, vol. 177, pp. 28-40, 2007.
[28] Z. Pawlak and A. Skowron, "Rough sets and boolean reasoning," *Inform. Sci.*, vol. 177, pp. 41-73, 2007.
[29] W. Pedrycz, "Relational and directional aspects in the construction of information granules," *IEEE Trans. Syst. Man. CY. A*, vol. 32, no. 5, pp. 605-614, 2002.
[30] W. Pedrycz and A. Bargiela, "Granular clustering: a granular signature of data," *IEEE Trans. Syst. Man. CY. B*, vol. 32, no. 2, pp. 212-224, 2002.
[31] L. Polkowski and A. Skowron, "Rough mereology: A new paradigm for approximate reasoning," *Intern. J. Approx. Reason.*, vol. 15, no. 4, pp. 333-365, 1996.
[32] L. Polkowski, A. Skowron and J. Zytkow, "Rough foundations for rough sets," in: T. Y. Lin and A. M. Wildberger (eds.), *Soft Computing: Rough Sets, Fuzzy Logic, Neural Networks, Uncertainty Management, Knowledge Discovery, Simulation Councils*, San Diego, CA, USA, pp. 55-58, 1995.
[33] Y. H. Qian, C. Y. Dang, and J. Y. Liang, "MGRS in incomplete information systems," *Proc. 2007 IEEE Intern. Conf. Granular Computing*, pp. 163-168, 2007.
[34] Y. H. Qian and J. Y. Liang, "Rough set method based on multi-granulations," *Proc. 5th IEEE Conf. Cognitive Informatics*, vol. I, pp. 297-304, 2006.

[35] Y. H. Qian, J. Y. Liang, and C. Y. Dang, "Converse approximation and rule extracting from decision tables in rough set theory," *Comput. Math. Appl.*, vol. 55. pp. 1754-1765, 2008.
[36] Y. H. Qian, J. Y. Liang, D. Y. Li, et al., "Measures for evaluating the decision performance of a decision table in rough set theory," *Inform. Sci.*, vol. 178, pp. 181-202, 2008.
[37] Y. H. Qian, C. Y. Dang, J. Y. Liang, et al., "On the evaluation of the decision performance of an incomplete decision table," *Data Knowledge Eng.*, vol. 65. no. 3, pp. 373-400, 2008.
[38] Y. H. Qian, J. Y. Liang, and C. Y. Dang, "Consistency measure, inclusion degree and fuzzy measure in decision tables," *Fuzzy Sets Syst.*, vol. 159. pp. 2353-2377, 2008.
[39] H. Rasiowa and V. Marek, "On reaching consensus by groups of intelligent agents", in: Z.W. Ras (Ed.) *Methodologies for Intelligent Systems*, NorthHolland, vol. 4, pp. 234-243, 1989.
[40] H. Rasiowa, "Mechanical proof systems for logic of reaching consensus by groups of intelligent agents", *J. Approx. Reasoni.*, vol. 5, pp. 415-432, 1991.
[41] C. Rauszer, "Knowledge representation systems for groups of agents", in: J. Wronski (Ed.) *Philosophical Logic in Poland*, Kluwer, Dordrecht, Netherlands, pp. 217-238, 1994.
[42] C. Rauszer and H. Swart, "Different approaches to knowledge, common knowledge and aumann's theorem", in: A. Lanux, H. Wansing (Eds.) *Knowledge and Belief in Philosophy and Artificial Intelligence*, Academie Verlag, Berlin, pp. 87-101, 1995.
[43] M. Rebolledo, "Rough intervals-enhancing intervals f or qualitative modeling of technical systems", *Artif. Intell.*, vol. 170, pp. 667-685, 2006.
[44] Q. Shen and A. Chouchoulas, "A rough-fuzzy approach for generating classification rules," *Pattern Recog.*, vol. 35, no. 11, pp. 2425-2438, 2002.
[45] A. Skowron and J. Stepaniuk, "Tolerance approximation spaces," *Fundamenta Informaticae.*, vol. 27, no. 2-3, pp. 245-253, 1996.
[46] D. Ślezak and W. Ziarko, "The investigation of the Bayesian rough set model," *J. Approx. Reason.*, vol. 40, pp. 81-91, 2005.
[47] R. Slowinski and D. Vanderpooten, "Similarity relation as a basis for rough approximations," in: P. Wang (Ed.), *Advances in Machine Intelligence and Soft Computing*, Duke University Press, Duke, NC, vol. 4, pp. 17-33, 1997.
[48] W. Swiniarski, Roman, and L. Hargis, "Rough sets as a front end of neural-networks texture classifiers," *Neurocomput.*, vol. 36, no. 1-4, pp. 85-102, 2001.
[49] W. Swiniarski and A. Skowron, "Rough set methods in feature selection and recognition," *Pattern Recog. Lett.*, vol. 24, no. 6, pp. 883-849, 2003.
[50] S. Tsumoto, "Automated extraction of hierarchical decision rules from clinical databases using rough set model," *Expert Syst. Appl.*, vol. 24, no. 2, pp. 189-197, 2003.
[51] G. Wang, H. Yu, and D. Yang, "Decision table reduction based on conditional information entropy," *Chinese J. Comp.*, vol. 25, no. 7, pp. 1-9, 2002.
[52] Y. Y. Yao, "Information granulation and rough set approximation," *Int. J. Intell. Syst.*, vol. 16, no. 1, pp. 87-104, 2001.
[53] Y. Y. Yao, "Granular computing basis issues and possible solutions," *Proc. Fifth Intern. Conf. Computing and Information.*, vol. I, pp. 186-189, 2000.
[54] Y. Y. Yao, "Perspectives of Granular Computing", *Proc. 2005 IEEE Intern. Conf. Granular Computing.*, vol. I, pp. 85-90, 2005.
[55] L. Zadeh, "Fuzzy logic equals computing with words," *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 2, pp. 103-111, 1996.
[56] L. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy Sets Syst.*, vol. 19, pp. 111-127, 1997.
[57] L. Zadeh, "A new direction in AI-Toward a computational theory of perceptions," *AI Mag.*, vol. 22, no. 1, pp. 73-84, 2001.
[58] A. Zeng, D. Pan, Q. L. Zheng and H. Peng, "Knowledge acquisition based on rough set theory and principal component analysis," *IEEE Intell. Syst.*, pp. 78-85, 2006.
[59] N. Zhong, J. Dong, and S. Ohsuga, "Rule discovery by soft induction techniques," *Neurocomput.*, vol. 36, no. 1-4, pp. 171-204, 2001.
[60] W. Ziarko, "Variable precision rough sets model", *J. Comput. Syst. Sci.*, vol. 46, no. 1, pp. 39-59, 1993.