Knowledge-Based Systems xxx (2011) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Distance: A more comprehensible perspective for measures in rough set theory

Jiye Liang, Ru Li, Yuhua Qian *

School of Computer and Information Technology, Shanxi University, Taiyuan 030006, Shanxi, China
Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Taiyuan 030006, Shanxi, China

ARTICLE INFO

Article history:
Received 6 October 2010
Received in revised form 26 September 2011
Accepted 3 November 2011
Available online xxxx

Keywords:
Rough set theory
Set distance
Partition distance
Measures

ABSTRACT

Distance provides a comprehensible perspective for characterizing the difference between two objects in a metric space. There are many measures which have been proposed and applied for various targets in rough set theory. In this study, through introducing set distance and partition distance to rough set theory, we investigate how to understand measures from rough set theory in the viewpoint of distance, which are inclusion degree, accuracy measure, rough measure, approximation quality, fuzziness measure, three decision evaluation criteria, information measure and information granularity. Moreover, a rough set framework based on the set distance is also a very interesting perspective for understanding rough set approximation. From the view of distance, these results look forward to providing a more comprehensible perspective for measures in rough set theory.

1. Introduction

Rough set theory proposed by Pawlak in [17] is a relatively new soft computing tool for the analysis of a vague description of an object, and has become a popular mathematical framework for pattern recognition, image processing, feature selection, neuro computing, conflict analysis, decision support, data mining and knowledge discovery from large data sets [1–3,13,23,30,39,42]. Rough-set-based data analysis starts from a data table, called information tables. The information tables contain data about objects of interest, characterized by a finite set of attributes. It is often interesting to discover some dependency relationships (patterns). An information table where condition attributes and decision attributes are distinguished is called a decision table. From a decision table one can induce some patterns in form of "*if*..., *then*..." decision rules [5,6,19,30]. More exactly, the decision rules say that if condition attributes have given values, then decision attributes have other given values.

To date, many measures for uncertainty have been proposed in rough set theory. As follows, for our further development, we briefly review several important measures. The concept of inclusion degree has been introduced into rough set theory, which is derived from the including measure among sets. Several authors have established several important relationships between inclusion degree and measures of rough set data analysis [27,40]. In rough set theory, as three classical measures, approximation accuracy, rough measure and approximate quality can be used to assess the roughness of a rough set and a rough classification [7,17]. For any object on a given universe, the membership function of the object in a rough set can be derived by the inclusion degree between the equivalence class including itself and a target concept, which can construct a fuzzy set on the universe. Several authors have studied the fuzziness of a rough set from various viewpoints [21,41]. In recent years, how to evaluate the decision performance of a decision rule and a decision-rule set has become a very important issue in rough set theory. There are two classical measures such as certainty measure and coverage measure [17]. In order to assess the decision performance of a decision table, Qian et al. [20] proposed three evaluation parameters α, β and γ which are used to calculate the entire certainty, the entire consistency and the entire support of all decision rules from a given decision table. However, each of the above measures is defined by different forms, which is hard to understand their sematic meanings. In other word, the uniform characterization of these measures is desirable. As we know, the concept of distance is a main approach to understand the difference between two objects in algebra, geometry, set theory, coding theory and many other areas. Hence, in this study, we aims to propose the concept of set distance to characterize and redefine each of these measures in order to more easily comprehend their meanings. It is exciting that Pawlak's rough set framework can be reconfigured using the set distance. This idea also can be used to redefined the variable precision rough set model proposed by Ziarko [46]. These results will be very helpful for us to understand the essence of rough set approximation. That is to say, it is a more comprehensible perspective for measures in rough set theory.

* Corresponding author at: School of Computer and Information Technology, Shanxi University, Taiyuan 030006, Shanxi, China. Tel./fax: +86 0351 7018176.
E-mail addresses: ljy@sxu.edu.cn (J. Liang), liru@sxu.edu.cn (R. Li), jinchengqyh@126.com (Y. Qian).

0950-7051/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2011.11.003

In addition, information entropy and information granularity are two main approaches to characterizing the uncertainty of an information system [14,18,28,43,45]. In recent years, several various forms of information entropy and information granularity have been given in [13,14,28,43,45]. It is deserved to point out that when the information granularity (or information entropy) of one equivalence partition is equal to that of the other equivalence partition, these two equivalence partitions have the same uncertainty. Nevertheless, it does not mean that these two equivalence partitions are equivalent. That is to say, information entropy and information granularity cannot characterize the difference between any two equivalence partitions in a given information table. In fact, we often need to distinguish two equivalence partitions for uncertain data processing in some practical applications. To date, how to measure the difference between equivalence partitions has not been reported. To further investigate uncertainty theory in the framework of rough set theory, for this consideration, we will propose the concept of partition distance to calculate the difference between two partitions on the same universe in this paper. In particular, we also reveal the essence of definitions of information entropy and information granularity from the viewpoint of partition distance.

The rest of this paper is organized as follows. Some preliminary concepts in rough set theory are briefly recalled in Section 2. In Section 3, we introduce the concept of set distance to characterize several important measures, which are inclusion degree, accuracy measure, rough measure, approximation quality, several decision evaluation parameters and the fuzziness measures of rough sets. In addition, we employ the set distance for reconfiguring the rough set framework and the variable precision rough set model. In Section 4, we first define the concept of partition distance to calculate the difference between two partitions on the same universe, then employ the partition distance to understand information entropy and information granularity from the viewpoint of distance. Section 6 concludes this paper with some remarks and discussions.

2. Preliminary knowledge in rough sets

In this section, we review some basic concepts such as indiscernibility relation, partition, partial relation of knowledge and decision tables in rough set theory.

An information table (sometimes called a data table, an attribute-value system, a knowledge representation system, etc.), as a basic concept in rough set theory, provides a convenient framework for the representation of objects in terms of their attribute values. An information table S is a pair (U,A), where U is a non-empty, finite set of objects and is called the universe and A is a non-empty, finite set of attributes. For each $a \in A$, a mapping $a: U \to V_a$ is determined by a given decision table, where V_a is the domain of a.

Each non-empty subset $B \subseteq A$ determines an indiscernibility relation in the following way,

$$R_B = \{(x,y) \in U \times U | a(x) = a(y), \forall a \in B\}.$$

The relation R_B partitions U into some equivalence classes given by

$$U/R_B = \{[x]_B | x \in U\}, \text{ just } U/B,$$

where $[x]_B$ denotes the equivalence class determined by x with respect to B, i.e.,

$$[x]_B = \{y \in U | (x,y) \in R_B\}.$$

Given an equivalence relation R on the universe U and a subset $X \subseteq U$. One can define a lower approximation of X and an upper approximation of X by

$$\underline{R}X = \{x \in U | [x]_R \subseteq X\}$$

and

$$\overline{R}X = \{x \in U | [x]_R \cap X \neq \emptyset\},$$

respectively [15]. The ordered pair $(\underline{R}X, \overline{R}X)$ is called a rough set of X with respect to R.

We define a partial relation $\preceq$ on the family $\{U/B | B \subseteq A\}$ as follows: $U/P \preceq U/Q$ (or $U/Q \succeq U/P$) if and only if, for every $P_i \in U/P$, there exists $Q_j \in U/Q$ such that $P_i \subseteq Q_j$, where $U/P = \{P_1, P_2, \ldots, P_m\}$ and $U/Q = \{Q_1, Q_2, \ldots, Q_n\}$ are partitions induced by $P, Q \subseteq A$, respectively. In this case, we say that Q is coarser than P, or P is finer than Q. If $U/P \preceq U/Q$ and $U/P \neq U/Q$, we say Q is strictly coarser than P (or P is strictly finer than Q), denoted by $U/P \prec U/Q$ (or $U/Q \succ U/P$).

It is clear that $U/P \prec U/Q$ if and only if, for every $X \in U/P$, there exists $Y \in U/Q$ such that $X \subseteq Y$, and there exist $X_0 \in U/P$, $Y_0 \in U/Q$ such that $X_0 \subset Y_0$.

A decision table is an information table $S = (U, C \cup D)$ with $C \cap D = \emptyset$, where an element of C is called a condition attribute, C is called a condition attribute set, an element of D is called a decision attribute, and D is called a decision attribute set. One can extract certain decision rules from a consistent decision table and uncertain decision rules from an inconsistent decision table.

3. Set distance and some measures in rough sets

The concept of set closeness between two classical sets is used to measure the degree of the sameness between sets. Let X and Y be two finite sets, the measure is defined by $H(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}$ $(X \cup Y \neq \emptyset)$. Obviously, the formula $1 - H(X,Y) = 1 - \frac{|X \cap Y|}{|X \cup Y|}$ can characterize the difference between two finite classical sets. In a broad sense, this measure can be regarded as a generalized distance [22]. Using the measure, one can obtain the following distance between two finite classical sets.

Definition 1. Let X, Y are two finite sets. The distance between X and Y is defined as

$$d(X,Y) = 1 - \frac{|X \cap Y|}{|X \cup Y|}, \tag{1}$$

where $X \cup Y \neq \emptyset$.

From the definition of the distance, one can easily obtain the following property.

Property 1. *The distance d satisfies the following properties:*

(1) $d(X,Y) \geqslant 0$;
(2) $d(X,Y) = d(Y,X)$;
(3) $d(X,Y) + d(Y,Z) \geqslant d(X,Z)$.

Proof. The three properties will be proved as follows.

(1) Obviously, $|X \cup Y| \geqslant |X \cap Y|$. Thus we have that,

$$d(X,\ Y) = \frac{|X \cup Y| - |X \cap Y|}{|X \cup Y|} \geqslant 0.$$

(2) It is easy to know that $|X \cup Y| = |Y \cup X|$ and $|X \cap Y| = |Y \cap X|$. Therefore,

$$d(X,\ Y) = \frac{|X \cup Y| - |X \cap Y|}{|X \cup Y|} = \frac{|Y \cup X| - |Y \cap X|}{|Y \cup X|} = d(Y,\ X).$$

(3) Given any a, b and c, and let $0 < b \leqslant a$, $c \geqslant 0$. From $\frac{b+c}{a+c} - \frac{b}{a} = \frac{c(a-b)}{a(a+c)} \geqslant 0$, it follows that $\frac{b}{a} \leqslant \frac{b+c}{a+c}$. Hence,

Please cite this article in press as: J. Liang et al., Distance: A more comprehensible perspective for measures in rough set theory, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.11.003

4

$$\begin{aligned}
&d(X,Y)+d(Y,Z)-d(X,Z)\\
&=1-\frac{|X\cap Y|}{|X\cup Y|}+1-\frac{|Y\cap Z|}{|Y\cup Z|}-1\\
&\quad+\frac{|X\cap Z|}{|X\cup Z|}=1-\frac{|X\cap Y|}{|X\cup Y|}-\frac{|Y\cap Z|}{|Y\cup Z|}+\frac{|X\cap Z|}{|X\cup Z|}\\
&\geqslant 1-\frac{|X\cap Y|+|Z|-|Z\cap(X\cup Y)|}{|X\cup Y\cup Z|}\\
&\quad-\frac{|Y\cap Z|+|X|-|X\cap(Y\cup Z)|}{|X\cup Y\cup Z|}+\frac{|X\cap Z|}{|X\cup Y\cup Z|}\\
&=1-\frac{|X\cap Y|+|Z|-(|X\cap Z|+|Y\cap Z|-|X\cap Y\cap Z|)}{|X\cup Y\cup Z|}\\
&\quad-\frac{|Y\cap Z|+|X|-(|X\cap Y|+|X\cap Z|-|X\cap Y\cap Z|)}{|X\cup Y\cup Z|}\\
&\quad+\frac{|X\cap Z|}{|X\cup Y\cup Z|}\\
&=1-\frac{|X|+|Z|-|X\cap Z|}{|X\cup Y\cup Z|}+\frac{2(|X\cap Z|-|X\cap Y\cap Z|)}{|X\cup Y\cup Z|}\\
&\geqslant 0
\end{aligned}$$

Therefore, $d(X,Y)+d(Y,Z)\geqslant d(X,Z)$. □

In this section, we establish the relationship between the set distance and each of several measures in rough set theory.

3.1. Set distance and inclusion degree

An approximate mereological calculus called rough mereology (i.e., theory of rough inclusions) has been proposed as a formal treatment of the hierarchy of relations of being a part in a degree. The degree of inclusion is a particular case of inclusion in a degree (rough inclusion) basic for rough mereology. The concept of inclusion degree based on partial relation was proposed in [46] for approximate reasoning in rough set theory. In the literature [40], Xu and Liang presented three types of inclusion degrees (I_0, I_1 and I_2), which have been successfully applied for characterizing the measures from rough set theory. In the following, we discuss the relationship between these three inclusion degrees and the set distance.

A partial order on a set L is a binary relation $\preceq$ with the following properties: $x\preceq x$ (reflexive), $x\preceq y$ and $y\preceq x$ imply $x=y$ (antisymmetric), and $x\preceq y$ and $y\preceq z$ imply $x\preceq z$ (transitive) [40].

Proposition 1. *Let U be a finite set, $F=\{X|X\subseteq U\}$ and $\subseteq$ a partial relation on F. For $\forall X,Y\in F$, one define an inclusion degree as*

$$I_0(Y/X)=\frac{|Y\cap X|}{|X|}. \tag{2}$$

It is easy to see that I_0 can be induced to a set distance

$I_0(Y/X)=d(X,\ X-Y)$.

Proposition 2. *Let $Y=\{Y_1,Y_2,\ldots,Y_n\}$ be a classification of U, $F=\{\{F_1,F_2,\ldots,F_n\}|F_i\subseteq Y_i,\ i=1,2,\ldots,n\}$, $X=\{X_1,X_2,\ldots,X_n\}\in F$ and $Z=\{Z_1,Z_2,\ldots,Z_n\}\in F$. One define another inclusion degree as*

$$I_1(X/Z)=\frac{\left|(\bigcup_{i=1}^n X_i)\cap(\bigcup_{i=1}^n Z_i)\right|}{\left|(\bigcup_{i=1}^n Z_i)\right|}. \tag{3}$$

From the formula, one can get the form of set distance of this inclusion degree, which is

$$I_1(X/Z)=d\Big(\Big(\bigcup\nolimits_{i=1}^n Z_i\Big),\Big(\bigcup\nolimits_{i=1}^n Z_i\Big)-\Big(\bigcup\nolimits_{i=1}^n X_i\Big)\Big).$$

Theorem 1. *Let $F=\{W|W\subseteq U\}$ and $P,Q\subseteq W$. Let $Y=\{Y_1,Y_2,\ldots,Y_n\}$ be a classification of U, $F'=\{\{F'_1,F'_2,\ldots,F'_n\}|F'_i\subseteq Y_i, i=1,2,\ldots,n\}$, $X=\{X_1,X_2,\ldots,X_n\}\in F'$ and $Z=\{Z_1,Z_2,\ldots,Z_n\}\in F'$. If $P=\bigcup_{i=1}^n X_i$ and $Q=\bigcup_{i=1}^n Y_i$, then $I_1(X/Z)$ is a special case of $I_0(P/Q)$.*

Proof. From the existing condition $P=\bigcup_{i=1}^n X_i$ and $Q=\bigcup_{i=1}^n Y_i$.

$$I_1(X/Z)=\frac{\left|(\bigcup_{i=1}^n X_i)\cap(\bigcup_{i=1}^n Z_i)\right|}{\left|(\bigcup_{i=1}^n Z_i)\right|}=\frac{|P\cap Q|}{|Q|}=I_0(P/Q).$$

Therefore, $I_1(X/Z)$ is a special case of $I_0(P/Q)$. □

From the above three propositions, it can been that three types of inclusion degrees can be all induced to be the set distance.

3.2. Set distance, accuracy measure, rough measure and approximation quality

As three classical measures, accuracy measure, rough measure and approximation quality are three important measures in rough set theory [17]. In this subsection, we investigate how to induce these kinds of measures to the set distance.

Proposition 3. *Let $S=(U,A)$ be an information table, $P\subseteq A$ and $X\subseteq U$. The accuracy measure of rough set X with respect to P [17] is defined as*

$$\alpha_P(X)=\frac{|\underline{P}X|}{|\overline{P}X|}, \tag{4}$$

where $X\neq\emptyset$.

It is easy to show that

$$\alpha_P(X)=\frac{|\underline{P}X\cap\overline{P}X|}{|\overline{P}X|}=d(\overline{P}X,\ \overline{P}X-\underline{P}X).$$

This shows that this measure can be induced to a set distance.

Proposition 4. *Let $S=(U,A)$ be an information table, $P\subseteq A$ and $Y=\{Y_1,Y_2,\ldots,Y_n\}$. By P-lower and P-upper approximation of Y in S we mean sets $\underline{P}Y=\{\underline{P}Y_1,\underline{P}Y_2,\ldots,\underline{P}Y_n\}$ and $\overline{P}Y=\{\overline{P}Y_1,\overline{P}Y_2,\ldots,\overline{P}Y_n\}$, respectively. The approximation quality of the classification Y with respect to P [17] is defined as*

$$r_P(Y)=\frac{\sum_{i=1}^n|\underline{P}Y_i|}{|U|}. \tag{5}$$

From the formula and Proposition 4, one can know that the measure

$$r_P(Y)=\frac{|U\cap\bigcup_{i=1}^n\underline{P}Y_i|}{|U|}=d\Big(U,\ U-\bigcup\nolimits_{i=1}^n\underline{P}Y_i\Big)$$

is also a set distance.

Proposition 5. *Let $S=(U,A)$ be an information table, $P\subseteq A$ and $X\subseteq U$. The rough measure of rough set X with respect to P [17] is defined as*

$$\rho_P(X) = 1 - \alpha_P(X) = 1 - \frac{|\underline{P}X|}{|\overline{P}X|}, \tag{6}$$

where $X \neq \emptyset$.

It is easy to show that

$$\rho_P(X) = d(\underline{P}X,\ \overline{P}X).$$

It implies that this measure also can be induced to a set distance.

The above three examples show approximation accuracy, rough measure and approximation quality can be all characterized by the set distance. In addition, from Proposition 6, it is easy to know that the measure of dependency between two attributes subsets also can be induced to a set distance.

3.3. Set distance and fuzziness measure of a rough set

In this subsection, we will research the set distance characterization of the fuzziness measures of a rough set and a rough decision.

Proposition 6. *Let* $S = (U, A)$ *be an information table and* $X \subseteq U$. *For any object* $x \in U$, *the membership function of* x *in* X *is defined as*

$$\mu_X^A(x) = \frac{|X \cap [x]_A|}{|[x]_A|}, \tag{7}$$

where $\mu_X^A(x) (0 \leqslant \mu_X^A(x) \leqslant 1)$ *represents a fuzzy concept* [36].

Obviously, the membership function can be redefined by the following set distance

$$\mu_X^A(x) = d([x]_A,\ [x]_A - X).$$

It can construct a fuzzy set $F_X^A = \{(x, \mu_X(x)) | x \in U\}$ on the universe U.

Proposition 7. *Let* $S = (U, A)$ *be an information table and* $X \subseteq U$. *A fuzziness measure of the rough set* X *is defined as* [21]

$$E\left(F_X^A\right) = \sum_{i=1}^{|U|} \mu_X^A(x)\left(1 - \mu_X^A(x)\right). \tag{8}$$

Through using the result of Proposition 6, it is obvious that

$$E\left(F_X^A\right) = \sum_{i=1}^{|U|} d([x]_A,\ [x]_A - X)(1 - d([x]_A,\ [x]_A - X)).$$

That is to say, the fuzziness measure also can be characterized by the set distance.

Proposition 8. *Let* $S = (U, A)$ *be an information table and* $U/D = \{Y_1, Y_2, \ldots, Y_n\}$ *a target decision. For any* $x \in U$, *the membership function of* x *in* D *is defined as* [13]

$$\mu_D(x) = \frac{|Y_j \cap [x]_A|}{|[x]_A|}, \quad x \in Y_j, \tag{9}$$

where $\mu_D(x)$ $(0 \leqslant \mu_D(x) \leqslant 1)$ *represents a fuzzy concept.*

Similar to Proposition 6, the membership function can be induced to the set distance

$$\mu_D(x) = d([x]_A,\ [x]_A - Y_j),\ x \in Y_j.$$

It can construct a fuzzy set $F_D^A = \{(x, \mu_D(x)) | x \in U\}$ on the universe U. Based on this membership function, one can construct a fuzziness measure of a rough decision.

Proposition 9. *Let* $S = (U, A)$ *be an information table and* $U/D = \{Y_1, Y_2, \ldots, Y_n\}$ *a target decision. A fuzziness measure of a rough decision is defined as* [21]

$$E\left(F_D^A\right) = \sum_{i=1}^{|U|} \mu_D(x_i)(1 - \mu_D(x_i)). \tag{10}$$

It can be depicted by the set distance

$$E\left(F_D^A\right) = \sum_{j=1}^{n} d([x]_A,\ [x]_A - Y_j)(1 - d([x]_A,\ [x]_A - Y_j)).$$

From the above four propositions, one know that membership functions and fuzziness measures can be induced to be a set distance in rough set theory. These cases will be helpful for understanding the fuzziness of a rough set and a rough decision by using set distance.

3.4. Set distance and decision performance evaluation

In recent years, how to evaluate the decision performance of a decision rule and a decision-rule set has become a very important issue in rough set theory. Firstly, we concern on two classical evaluation measures, which are certainty measure and coverage measure.

Let $S = (U, C \cup D)$ be a decision table, $X_i \in U/C$, $Y_j \in U/D$ and $X_i \cap Y_j \neq \emptyset$. By des (X_i) and des (Y_j), we denote the descriptions of the equivalence classes X_i and Y_j in the decision table S [17]. A decision rule is formally defined ~~[2]~~

$$Z_{ij}:\ des(X_i) \rightarrow des(Y_j).$$

Proposition 10. *Let* $S = (U, C \cup D)$ *be a decision table,* $X_i \in U/C$, $Y_j \in U/D$ *and* $X_i \cap Y_j \neq \emptyset$. *Certainty measure (also called resolution) of the rule* Z_{ij} *is defined as* [37]

$$\alpha_{X_i}(Y_j) = \frac{|Y_j \cap X_i|}{|X_i|}. \tag{11}$$

From the definition of certainty measure, it is easy to see that the formula (11) can be depicted by a set distance

$$\alpha_{X_i}(Y_j) = d(X_i,\ X_i - Y_j).$$

Proposition 11. *Let* $S = (U, C \cup D)$ *be a decision table,* $X_i \in U/C$, $Y_j \in U/D$ *and* $X_i \cap Y_j \neq \emptyset$. *Coverage measure (also called completeness) of the rule* Z_{ij} *is defined as* [37]

$$\beta_{X_i}(Y_j) = \frac{|Y_j \cap X_i|}{|Y_j|}. \tag{12}$$

From the denotation of coverage measure, it can be seen that the formula (12) also can be characterized by a set distance

$$\beta_{X_i}(Y_j) = d(Y_j,\ Y_j - X_i).$$

Similar to the support measure of a decision rule $s(Z_{ij}) = \frac{|Y_j \cap X_i|}{|U|}$ is also induced to the set distance $D(U, U - (Y_j \cap X_i))$.

In order to assess the decision performance of a decision table, Qian et al. [20] proposed three evaluation parameters α, β and γ, which are used to calculate the entire certainty, the entire consistency and the entire support of decision rules based on elementary sets from a given decision table.

Proposition 12. *Let $S=(U,C\cup D)$ be a decision table, and $RULE=\{Z_{ij}: des(X_i)\rightarrow des(Y_j),\ X_i\in U/C, Y_j\in U/D\}$. Certainty measure α of S is defined as*

$$\alpha(S)=\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|X_i\cap Y_j|^2}{|U||X_i|},\tag{13}$$

where $s(Z_{ij})$ and $\mu(Z_{ij})$ are the certainty measure and support measure of the rule Z_{ij}, respectively.

Through using Eq. (11), it easily follows that

$$\alpha(S)=\sum_{i=1}^{m}\sum_{j=1}^{n}d(X_i,\ X_i-Y_j)d(U,\ U-(X_i\cap Y_j)).$$

That is to say, the certainty measure can be defined by two set distances. Similarly, one can apply the set distance for depicting the following two evaluation criteria.

Proposition 13. *Let $S=(U,C\cup D)$ be a decision table, and $RULE=\{Z_{ij}: des(X_i)\rightarrow des(Y_j),\ X_i\in U/C, Y_j\in U/D\}$. Consistency measure β of S is defined as*

$$\beta(S)=\sum_{i=1}^{m}\frac{|X_i|}{|U|}\left[1-\frac{4}{|X_i|}\sum_{j=1}^{N_i}|X_i\cap Y_j|\mu(Z_{ij})(1-\mu(Z_{ij}))\right],\tag{14}$$

where N_i is the number of decision rules by the condition class X_i and $\mu(Z_{ij})$ is the certainty measure of the rule Z_{ij}.

If adopting the interpretation of set distance, one can obtain the following denotation.

$$\beta(S)=1-\frac{4}{|U|}\sum_{i=1}^{m}\sum_{j=1}^{N_i}|X_i\cap Y_j|d(X_i,\ X_i-Y_j)(1-d(X_i,\ X_i-Y_j)).$$

Proposition 14. *Let $S=(U,C\cup D)$ be a decision table, and $RULE=\{Z_{ij}: des(X_i)\rightarrow des(Y_j),\ X_i\in U/C, Y_j\in U/D\}$. Support measure γ of S is defined as*

$$\gamma(S)=\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{|X_i\cap Y_j|^2}{|U|^2}.\tag{15}$$

From the definition of set distance, it is clear that

$$\gamma(S)=\sum_{i=1}^{m}\sum_{j=1}^{n}d^2(U,\ U-(X_i\cap Y_j)).$$

From the above these propositions, we also can use the set distance to characterizing those evaluation measures of decision performance in the context of incomplete decision tables.

3.5. Rough set framework based on set distance

In rough set theory, the characterization of a target concept is approximated by the lower approximation and the upper approximation. In order to better comprehend the idea from rough set theory, in this subsection, we will apply the set distance for redefining the concept of a rough set.

Given an equivalence relation R on the universe U and a subset $X\subseteq U$. One can define a lower approximation of X and an upper approximation of X by

$$\underline{R}X=\{x\in U|d([x]_R,\ X\cap[x]_R)=0\},\tag{16}$$
$$\overline{R}X=\{x\in U|0\leqslant d([x]_R,\ X\cap[x]_R)<1\}.\tag{17}$$

This definition of the rough set, in fact, is equivalent to Pawlak's rough set. It can be understood from the following analysis.

In Pawlak's rough set, when $[x]_R\subseteq X$, the object x can be putted into the lower approximation of X, and when $[x]_R\cap X\neq Ø$, the object x can be putted into the upper approximation of X. From the condition $[x]_R\subseteq X$, we have that

$$[x]_R\subseteq X\Longleftrightarrow\frac{|(X\cap[x]_R)\cap[x]_R|}{|(X\cap[x]_R)\cup[x]_R|}=1\Longleftrightarrow 1-\frac{|(X\cap[x]_R)\cap[x]_R|}{|(X\cap[x]_R)\cup[x]_R|}=0\Longleftrightarrow d([x]_R,\ X\cap[x]_R)=0.$$

From $[x]_R\cap X\neq Ø$, we obtain that

$$[x]_R\cap X\neq Ø\Longleftrightarrow 0<\frac{|(X\cap[x]_R)\cap[x]_R|}{|(X\cap[x]_R)\cup[x]_R|}\leqslant 1\Longleftrightarrow 0\leqslant 1-\frac{|(X\cap[x]_R)\cap[x]_R|}{|(X\cap[x]_R)\cup[x]_R|}<1\Longleftrightarrow 0\leqslant d([x]_R,\ X\cap[x]_R)<1.$$

Similar to these two denotations, we come to the definitions of negative region and boundary region of a rough set as follows

$$Neg_RX=\{x\in U|d([x]_R,\ X\cap[x]_R)=1\},\tag{18}$$
$$Bn_RX=\{x\in U|0<d([x]_R,\ X\cap[x]_R)<1\}.\tag{19}$$

These two Eqs. (18) and (19) also can be similarly proved according to the analysis about Eqs. (16) and (17).

From Eqs. (16)–(19), it can be seen that the characterization of a rough set only depends on the set distance between $[x]_R$ and $X\cap[x]_R$. If the distance between the equivalence class $[x]_R$ and $X\cap[x]_R$ achieve the minimum value zero, then the equivalence class must be included in the lower approximation of the target concept. If the distance between them equal the maximum value one, then the equivalence class must belong to the negative region. The rest equivalence classes lie in the boundary region of the rough set.

Example 1. Let $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10},x_{11},x_{12}\}$ and R be an equivalence relation induced by $U/R=\{\{x_1,x_3\},\{x_2,x_4,x_5,x_6\},\{x_7,x_8,x_9,x_{10}\},\{x_{11},x_{12}\}\}$. We have that

$$[x_1]_R=[x_3]_R=\{x_1,x_3\},$$
$$[x_2]_R=[x_4]_R=[x_5]_R=[x_6]_R=\{x_2,x_4,x_5,x_6\},$$
$$[x_7]_R=[x_8]_R=[x_9]_R=[x_{10}]_R=\{x_7,x_8,x_9,x_{10}\}$$

and

$$[x_{11}]_R=[x_{12}]_R=\{x_{11},x_{12}\}.$$

Given a set $X=\{x_1,x_2,x_3,x_5,x_6,x_7,x_8,x_{11},x_{12}\}$. For Pawlak's rough set model, we can obtain the lower approximation and upper approximation of X

$$\underline{R}X=\{x\in U|[x]_R\subseteq X\}=\{x_1,x_3\},$$
$$\overline{R}X=\{x\in U|[x]_R\cap X\neq\emptyset\}=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\},$$
$$Neg_RX=U-\overline{R}X=\{x_{11},x_{12}\}$$

and

$$Bn_RX=\overline{R}X-\underline{R}X=\{x_2,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}.$$

For the above rough set framework based on the set distance, we can calculate the following results

$$d([x_1]_R,\ X\cap[x_1]_R)=d([x_3]_R,\ X\cap[x_3]_R)=1-\frac{|(X\cap[x_1]_R)\cap[x_1]_R|}{|(X\cap[x_1]_R)\cup[x_1]_R|}=0,$$

$$\begin{aligned} d([x_2]_R,\ X\cap[x_2]_R) &= d([x_4]_R,\ X\cap[x_4]_R) = d([x_5]_R,\ X\cap[x_5]_R) \\ &= d([x_6]_R,\ X\cap[x_6]_R) = 1-\frac{|(X\cap[x_2]_R)\cap[x_2]_R|}{|(X\cap[x_2]_R)\cup[x_2]_R|} \\ &= 1-\frac{3}{4} = 0.25, \end{aligned}$$

$$\begin{aligned} d([x_7]_R,\ X\cap[x_7]_R) &= d([x_8]_R,\ X\cap[x_8]_R) = d([x_9]_R,\ X\cap[x_9]_R) \\ &= d([x_{10}]_R,\ X\cap[x_{10}]_R) = 1-\frac{|(X\cap[x_7]_R)\cap[x_7]_R|}{|(X\cap[x_7]_R)\cup[x_7]_R|} \\ &= 1-\frac{2}{4} = 0.5 \end{aligned}$$

and

$$\begin{aligned} d([x_{11}]_R,\ X\cap[x_{11}]_R) &= d([x_{12}]_R,\ X\cap[x_{12}]_R) = 1-\frac{|(X\cap[x_{11}]_R)\cap[x_{11}]_R|}{|(X\cap[x_{11}]_R)\cup[x_{11}]_R|} \\ &= 1-0 = 1. \end{aligned}$$

Furthermore, we have that

$$\begin{aligned} &\underline{R}X = \{x\in U \mid d([x]_R,\ X\cap[x]_R) = 0\} = \{x_1, x_3\}, \\ &\overline{R}X = \{x\in U \mid 0 \leqslant d([x]_R,\ X\cap[x]_R) < 1\} \\ &\quad = \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}, \\ &Neg_R X = \{x\in U \mid d([x]_R,\ X\cap[x]_R) = \{x_{11},x_{12}\} \end{aligned}$$

and

$$\begin{aligned} Bn_R X &= \{x\in U \mid 0 \leqslant d([x]_R,\ X\cap[x]_R) < 1\} \\ &= \{x_2,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}. \end{aligned}$$

From these above equations, it can be seen that the R-lower approximation, R-upper approximation, R-negative region and R-boundary region of X obtained by Pawlak's rough set model are the same as the ones achieved by the proposed rough set framework based on the set distance. However, the rough set framework based on the set distance seems more intuitive for understanding the meaning of rough set approximation. That is to say, the set distance-based rough set model gives a more comprehensible perspective.

According to the definition of set distance, one easily obtains the following theorem.

Theorem 2. *Let R be an equivalence relation on U and $X \subseteq U$ a subset. Then,*

(1) *X is a R-definable set iff $d(\underline{R}X, \overline{R}X) = 0$;*
(2) *X is a R-rough set iff $d(\underline{R}X, \overline{R}X) > 0$.*

Proof. Similar to the proofs about Eqs. (16)–(19), this theorem can be easily proved. □

With the introduction of rough inclusion, the standard approximation space can be generalized to variable precision approximations. In formulating the variable precision rough set model, Ziarko [46] used the relative degree of misclassification function c and the granule based definition of approximation. In the variable rough set framework, one needs to choose the threshold value β in the range $[0, 0.5]$. Given an equivalence relation R on the universe U and a subset $X \subseteq U$. In variable rough set theory, through using the set distance, a lower approximation of X and an upper approximation of X can be redefined by

$$\underline{R}_\beta X = \{x\in U \mid d([x]_R,\ X\cap[x]_R) \leqslant \beta\}, \tag{20}$$

$$\overline{R}_\beta X = \{x\in U \mid 0 \leqslant d([x]_R,\ X\cap[x]_R) < 1-\beta\}. \tag{21}$$

Similar to these two denotations, we come to the definitions of negative region and boundary region of a variable rough set as follows

$$Neg_\beta X = \{x\in U \mid d([x]_R,\ X\cap[x]_R) \geqslant 1-\beta\}, \tag{22}$$

$$Bn_\beta X = \{x\in U \mid \beta < d([x]_R,\ X\cap[x]_R) < 1-\beta\}. \tag{23}$$

When the threshold value β equals zero, a variable rough set will degenerate into the corresponding Pawlak's rough set. From Eqs. (21)–(24), it can be seen that like the standard rough set, the depiction of a variable rough set also depends on the set distance between $[x]_R$ and $X\cap[x]_R$. Clearly, the characterization of a rough set by the set distance will be very helpful for more easily understanding the essence and meaning of a rough set, which provides a more comprehensible perspective for measures from rough set theory.

Example 2 (*Continued* from Example 1). Let $\beta = 0.4$, for Ziarko's variable precise rough set model, we can compute the lower approximation and upper approximation of X are as follows

$$\begin{aligned} &\underline{R}_\beta X = \{x\in U \mid 1-\frac{[x]_R\cap X}{[x]_R} \leqslant 0.4\} = \{x_1,x_2,x_3\ x_4,x_5,x_6\}, \\ &\overline{R}_\beta X = \{x\in U \mid 1-\frac{[x]_R\cap X}{[x]_R} < 0.6\} = \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}, \\ &Neg_\beta X = U - \overline{R}X = \{x_{11},x_{12}\} \end{aligned}$$

and

$$Bn_\beta X = \overline{R}X - \underline{R}X = \{x_7,x_8,x_9,x_{10}\}.$$

Using the proposed rough set framework based on the set distance, we obtain its lower approximation and upper approximation

$$\begin{aligned} &\underline{R}_\beta X = \{x\in U \mid d([x]_R,\ X\cap[x]_R) \leqslant 0.4\} = \{x_1,x_2,x_3\ x_4,x_5,x_6\}, \\ &\overline{R}_\beta X = \{x\in U \mid 0 \leqslant d([x]_R,\ X\cap[x]_R) < 0.6\} \\ &\quad = \{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}, \\ &Neg_\beta X = \{x\in U \mid d([x]_R,\ X\cap[x]_R) \geqslant 0.6 = \{x_{11},x_{12}\}\} \end{aligned}$$

and

$$Bn_\beta X = \{x\in U \mid 0.4 < d([x]_R,\ X\cap[x]_R) < 0.6\} = \{x_7,x_8,x_9,x_{10}\}.$$

From these above computations, it is easy to see that the R-lower approximation, R-upper approximation, R-negative region and R-boundary region of X obtained by VPRS are the same as the ones achieved by the redefined VPRS based on the set distance, respectively. Like the set distance-based rough set model, the redefined VPRS play the same role for understanding the meaning of rough set approximation in VPRS, which also displays a more comprehensible perspective.

In the literature [4], Cornelis et al. gone one step further by introducing vague quantifiers like most and some into the model VPRS. In this way, an element x belongs to the lower approximation of X if most of the elements related to x are included in X. Likewise, an element belongs to the upper approximation of X if some of the elements related to x are included in X. In this approach, it is implicitly assumed that the approximations are fuzzy sets, i.e., mappings from X to $[0, 1]$, that evaluate to what degree the associated condition is fulfilled. The authors formally define the upper approximation and lower approximation of X by fixing a couple of fuzzy quantifiers, which are also constructed based on the inclusion degree. Hence, we also can employ the proposed set distance for redefining so-called vaguely quantified rough set framework proposed by Cornelis et al. Due to their similarity, we omit its form based on the set distance.

Please cite this article in press as: J. Liang et al., Distance: A more comprehensible perspective for measures in rough set theory, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.11.003

4. Partition distance and some measures in rough sets

In rough set theory, information entropy and knowledge granulation are two main approaches to measuring the uncertainty of a partition in knowledge bases (approximation spaces). If the knowledge granulation (or information entropy) of one partition is equal to that of the other partition, we say that these two partitions have the same uncertainty. However, it does not mean that these two partitions are equivalent. In other words, information entropy and knowledge granulation cannot characterize the difference between any two partitions in a knowledge base. In this section, we introduce a notion of partition distance to differentiate two given partitions and investigate some of its important properties.

For our further development, we give several representations and denotations. 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. In this paper, we denote an equivalence partition induced by U/R on U by $K(R)$. In fact, the partition can be formally defined as $K(R)=\{E_R(x)|x\in U\}$. Each equivalence class $E_R(x)(x\in U)$ may be viewed as an information granule consisting of indistinguishable elements [22].

4.1. Partition distance

To characterize the relationship among partitions, based on the view of set distance, we introduce an approach called partition distance for measuring the difference between two partitions on the same knowledge base in the following.

Definition 2. Let $K=(U,\mathbf{R})$ be a knowledge base, $P,\ Q\in\mathbf{R}$, $K(P)=\{[x_i]_P|x_i\in U\}$ and $K(Q)=\{[x_i]_Q|x_i\in U\}$. Partition distance between $K(P)$ and $K(Q)$ is defined as

$$D(K(P),K(Q))=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P\bigoplus[x_i]_Q|}{|U|},\tag{24}$$

where $|[x_i]_P\oplus[x_i]_Q|=|[x_i]_P\cup[x_i]_Q|-|[x_i]_P\cap[x_i]_Q|$.

Theorem 3 (Extremum). *Let $K(U,\mathbf{R})$ be a knowledge base, $K(P)$, $K(Q)$ two partitions on K. Then, $D(K(P),K(Q))$ achieves its minimum value $D(K(P),K(Q))=0$ if $K(P)=K(Q)$ and $D(K(P),K(Q))$ achieves its maximum value $D(K(P),K(Q))=1-\frac{1}{|U|}$ if $K(P)=\omega$ and $K(Q)=\delta$ ($K(P)=\delta$ and $K(Q)=\omega$).*

Proof. For $\forall P,Q\in\mathbf{R}$, one has that $1\leqslant|[x_i]_P\cap[x_i]_Q|\leqslant|U|$, $1\leqslant|[x_i]_P\cup[x_i]_Q|\leqslant|U|$. Therefore, for $\forall P,Q\in\mathbf{R}$,

$$0\leqslant|[x_i]_P\oplus[x_i]_Q|\leqslant|U|-1,\ \text{i.e.,}\ 0\leqslant\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P\oplus[x_i]_Q|}{|U|}\leqslant1-\frac{1}{|U|}.$$

If $K(P)=K(Q)$, then $[x_i]_P\cap[x_i]_Q=[x_i]_P$, $[x_i]_P\cup[x_i]_Q=[x_i]_P$, $i\leqslant|U|$. Hence, $D(K(P),K(Q))=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P\oplus[x_i]_Q|}{|U|}=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{0}{|U|}=0$, i.e., $D(K(P),K(Q))$ achieves its minimum value 0.

If $K(P)=\omega$ and $K(Q)=\delta$, then $[x_i]_P\cap[x_i]_Q=\{x_i\}$, $[x_i]_P\cup[x_i]_Q=U$, $i\leqslant|U|$. Hence, $D(K(P),K(Q))=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P\oplus[x_i]_Q|}{|U|}=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|U|-|x_i|}{|U|}$ $=1-\frac{1}{|U|}$, i.e., $D(K(P),K(Q))$ achieves its maximum value $1-\frac{1}{|U|}$. □

The partition distance represents the measure of difference between two partitions in the same knowledge base. Obviously, $0\leqslant D(K(P),K(Q))\leqslant1-\frac{1}{|U|}$.

Let $K(P)=\{[x_i]_P|x_i\in U\}$, $K(Q)=\{[x_i]_Q|x_i\in U\}$ and $K(R)=\{[x_i]_R|x_i\in U\}$ be three partitions on U. For $[x_i]_P\in K(P)$, $[x_i]_Q\in K(Q)$ and $[x_i]_R\in K(R)$, $x_i\in U$, we denote $[x_i]_{\{(P,Q,R)\}}=[x_i]_P\cup[x_i]_Q\cup[x_i]_R$. One can give a certain array of all elements in $[x_i]_{P\cup Q\cup R}$ and denote the array by $Array=(x_{i_1},x_{i_2},\ldots,x_{i_{|[x_i]_{P\cup Q\cup R}|}})$. Therefore, one can represent $[x_i]_P$ by the following array

$$x_{i_k}=\begin{cases}1 & \text{if } x_{i_k}\in[x_i]_P,\\ 0 & \text{else},\end{cases}$$

for $x_{i_k}\in Array, k\leqslant|[x_i]_{P\cup Q\cup R}|$.

Similarly, the expressions of $[x_i]_Q$ and $[x_i]_R$ also can be obtained. In fact, the expression of $Array$ is various, so the expression of $[x_i]_P$, $[x_i]_Q$ and $[x_i]_R$ should be also changed according to $Array$, respectively. This kind of representations about the equivalence classes are illustrated by the following proposition.

Example 3. Consider three equivalence classes $[x_i]_P=\{1,2,3\}$, $[x_i]_Q=\{2,3,4\}$ and $[x_i]_R=\{3,4,5\}$. Compute the expressions of $[x_i]_P$, $[x_i]_Q$ and $[x_i]_R$ through using the above method.

By computing, one has that $[x_i]_{\{(P,Q,R)\}}=[x_i]_P\cup[x_i]_Q\cup[x_i]_R=\{1,2,3,4,5\}$. Assume that $Array=(1,2,3,4,5)$.

For $[x_i]_P$, one can obtain that $[x_i]_P=(1,1,1,0,0)$. Similarly, it follows that $[x_i]_Q=(0,1,1,1,0)$ and $[x_i]_R=(0,0,1,1,1)$.

Let A, B, C be three classical sets, $Array=(t_1,t_2,\ldots,t_{|\{(A,B,C)\}|})$, $t_i\neq t_j$, t_i, $t_j\in A\cup B\cup C$. Hence, from the above expression method, one can get the array expressions of A, B and C as follows

$$A'=(a_1,a_2,\ldots,a_{|\{(A,B,C)\}|}),$$

$$B'=(b_1,b_2,\ldots,b_{|\{(A,B,C)\}|})\ \text{and}$$

$$C'=(c_1,c_2,\ldots,c_{|\{(A,B,C)\}|}).$$

Based on these denotations, we then measure the distance between two classical sets by the following formula

$$d'(A,B)=\sum_{i=1}^{|\{(A,B,C)\}|}(a_i\oplus b_i),a_i\in A',b_i\in B'.\tag{25}$$

Analogously, one has that $d'(B,C)=\sum_{i=1}^{|\{(A,B,C)\}|}(b_i\oplus c_i)$ and $d'(A,C)=\sum_{i=1}^{|\{(A,B,C)\}|}(a_i\oplus c_i)$.

From these denotations, we come to the following lemma.

Lemma 1. *Let A, B, C be three classical sets, then $d'(A,B)+d'(B,C)\geqslant d'(A,C)$.*

Proof. Suppose that $A'=\{a_1,a_2,\ldots,a_{|\{(A,B,C)\}|}\}$, $B'=\{b_1,b_2,\ldots,b_{|\{(A,B,C)\}|}\}$ and $C'=\{c_1,c_2,\ldots,c_{|\{(A,B,C)\}|}\}$. From $(a_i\oplus b_i)\oplus(b_i\oplus c_i)\geqslant(a_i\oplus c_i)$, it follows that

$$\begin{aligned}d'(A,B)+d'(B,C)&=\sum_{i=1}^{|\{(A,B,C)\}|}(a_i\oplus b_i)+\sum_{i=1}^{|\{(A,B,C)\}|}(b_i\oplus c_i)\\&=\sum_{i=1}^{|\{(A,B,C)\}|}((a_i\oplus b_i)\oplus(b_i\oplus c_i))\geqslant\sum_{i=1}^{|\{(A,B,C)\}|}(a_i\oplus c_i)\\&=d'(B,C).\end{aligned}$$

Similarly, $d'(A,B)+d'(A,C)\geqslant d'(B,C)$ and $d'(A,C)+d'(B,C)\geqslant d'(A,B)$. □

Theorem 4. *Let $\mathbf{K}(U)$ be the set of all partitions induced by U, then $(\mathbf{K}(U),D)$ is a distance space.*

Proof

(1) One can obtain easily that $D(K(P),K(Q))\geqslant0$ from Definition 2.
(2) It is obvious that $D(K(P),K(Q))=D(K(Q),K(P))$.

Please cite this article in press as: J. Liang et al., Distance: A more comprehensible perspective for measures in rough set theory, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.11.003

(3) For the proof of the triangle inequality principle, one only need to prove that $D(K(P),K(Q)) + D(K(P),K(R)) \geqslant D(K(Q),K(R))$, $K(P),K(Q),K(R) \in \mathbf{K}(U)$.

From Lemma 1, we know that for $x_i \in U$, $d'([x_i]_P, [x_i]_Q) + d'([x_i]_P, [x_i]_R) \geqslant d'([x_i]_Q, [x_i]_R)$. Hence,

$$
\begin{aligned}
&D(K(P),K(Q)) + D(K(P),K(R)) \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P \oplus [x_i]_Q|}{|U|} + \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P \oplus [x_i]_R|}{|U|} \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{d'([x_i]_P,[x_i]_Q)}{|U|} + \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{d'([x_i]_P,[x_i]_R)}{|U|} \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{1}{|U|}\left(d'([x_i]_P,[x_i]_Q) + d'([x_i]_P,[x_i]_R)\right) \\
&\geqslant \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{d'([x_i]_Q,[x_i]_R)}{|U|} \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}D(K(Q),K(R)).
\end{aligned}
$$

Therefore, $(\mathbf{K}(U),D)$ is a distance space. □

From the above theorem, one can draw a conclusion that the partition distance is also distance metric.

For further development, we give the following Lemma 2.

Lemma 2. *Let A, B, C be three classical sets with $A \subseteq B \subseteq C$ or $A \supseteq B \supseteq C$, then $d'(A,B) + d'(B,C) = d'(A,C)$.*

Proof. Suppose that $A' = \{a_1, a_2, \ldots, a_{|\{(A,B,C)\}|}\}$, $B' = \{b_1, b_2, \ldots, b_{|\{(A,B,C)\}|}\}$ and $C' = \{c_1, c_2, \ldots, c_{|\{(A,B,C)\}|}\}$. Let $A \supseteq B \supseteq C$, thus $A \cup B \cup C = A$ and $B \cup C = B$. Therefore,

$$
\begin{aligned}
d'(A,B) + d'(B,C) &= \sum_{i=1}^{|\{(A,B,C)\}|}(a_i \oplus b_i) + \sum_{i=1}^{|\{(A,B,C)\}|}(b_i \oplus c_i) \\
&= (|A \cup B| - |A \cap B|) + (|B \cup C| - |B \cap C|) \\
&= (|A| - |B|) + (|B| - |C|) = |A| - |C| \\
&= \sum_{i=1}^{|\{(A,B,C)\}|}(a_i \oplus c_i) = d'(A,C).
\end{aligned}
$$

For $A \subseteq B \subseteq C$, similarly, one can draw the same conclusion. □

By Definition 2 and Lemma 2, one can obtain the following theorem.

Theorem 5. *Let $K = (U, \mathbf{R})$ be a knowledge base, $P, Q, R \in \mathbf{R}$ and $K(P) \preceq K(Q) \preceq K(R)$ or $K(R) \preceq K(Q) \preceq K(P)$. Then, $D(K(P),K(R)) = D(K(P),K(Q)) + D(K(Q),K(R))$.*

Proof. For $K(P),K(Q),K(R) \in K$ and $K(P) \preceq K(Q) \preceq K(R)$, one can easily get that $[x_i]_P \subseteq [x_i]_Q \subseteq [xi]_R$, $x_i \in U$. Hence, it follows from Lemma 2 that

$$
\begin{aligned}
&D(K(P),K(Q)) + D(K(Q),K(R)) \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P \oplus [x_i]_Q|}{|U|} + \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_Q \oplus [x_i]_R|}{|U|} \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{d'([x_i]_P,[x_i]_Q)}{|U|} + \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{d'([x_i]_Q,[x_i]_R)}{|U|} \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{1}{|U|}\left(d'([x_i]_P,[x_i]_Q) + d'([x_i]_Q,[x_i]_R)\right) \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{d'([x_i]_P,[x_i]_R)}{|U|} = \frac{1}{|U|}\sum_{i=1}^{|U|}D(K(P),K(R)).
\end{aligned}
$$

For $K(R) \preceq K(Q) \preceq K(P)$, similarly, one can draw the same conclusion. □

Example 4. Let $U = \{x_1,x_2,x_3,x_4,x_5\}$ and $K(P)$, $K(Q)$, $K(R)$ be three partitions induced by equivalence relations P, Q, R on K, where $K(P) = \{\{x_1,x_2,x_3\}, \{x_1,x_2,x_3\}, \{x_1,x_2,x_3\}, \{x_4,x_5\}, \{x_4,x_5\}\}$, $K(Q) = \{\{x_1,x_2\},\{x_1,x_2\}, \{x_3\}, \{x_4,x_5\}, \{x_4,x_5\}\}$ and $K(R) = \{\{x_1,x_2\}, \{x_1,x_2\}, \{x_3\}, \{x_4\}, \{x_5\}\}$.

It is obvious that $K(R) \preceq K(Q) \preceq K(P)$. By computing the partition distances among them, one can obtain that

$$
\begin{aligned}
D(K(P),K(Q)) &= \frac{1}{5}\left[\frac{1}{5}(1+1+2+0+0)\right] = \frac{4}{25}, \\
D(K(Q),K(R)) &= \frac{1}{5}\left[\frac{1}{5}(0+0+0+1+1)\right] = \frac{2}{25} \text{ and} \\
D(K(P),K(R)) &= \frac{1}{5}\left[\frac{1}{5}(1+1+2+1+1)\right] = \frac{6}{25}.
\end{aligned}
$$

It is clear that $D(K(P),K(Q)) + D(K(Q),K(R)) = \frac{4}{25} + \frac{2}{25} = \frac{6}{25} = D(K(P),K(R))$.

As a result of the above discussions and analyses, we come to the following corollary.

Corollary 1. *Let $\mathbf{K}(U)$ be the set of all partitions induced by U and $K(P)$ a partition on $\mathbf{K}(U)$, then $D(K(P),K(\delta)) + D(K(P),K(\omega)) = 1 - \frac{1}{|U|}$.*

Proof. Since $K(\omega) \preceq K(P) \preceq K(\delta)$, one can obtain that

$$
\begin{aligned}
&D(K(P),K(\delta)) + D(K(P),K(\omega)) \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P| - 1}{|U|} + \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|U| - |[x_i]_P|}{|U|} \\
&= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P| - 1 + |U| - |[x_i]_P|}{|U|} = \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|U| - 1}{|U|} = 1 - \frac{1}{|U|}.
\end{aligned}
$$

Obviously, $D(K(P),K(\delta)) + D(K(P),K(\omega)) = 1 - \frac{1}{|U|}$. □

Hence, the partition distance can characterize the difference between two partitions in knowledge bases.

4.2. Partition distance and information measure in information tables

An information measure can calculate the information content of an information table [13,15]. Let $S = (U,A)$ be a complete information table, $P \subseteq A$, the information measure $I(P)$ (I is an information measure function) of $K(P)$ should satisfy [7].

(1) $I(P) \geqslant 0$;
(2) if $K(P) = K(Q)$, then $I(P) = I(Q)$; and
(3) if $K(P) \prec K(Q)$, then $I(P) > I(Q)$.

For example, each of Shannon's information entropy [28] and Liang's information entropy [14] is an information measure which is used to measure the information content of a complete information table.

In what follows, we establish the relationship between the partition distance and information measure in information tables.

Theorem 6. *$D(K(P),K(\delta))$ is an information measure.*

Proof. Let $K(\delta)=\{[x_i]_\delta|[x_i]_\delta=U, x_i\in U\}$ and $K(P)=\{[x_i]_P|x_i\in U\}$.

(1) This distance D is clearly non-negative.
(2) If $K(P)=K(Q)$, then $D(K(P),K(\delta))=D(K(Q),K(\delta))$.
(3) We prove that if $K(P)\prec K(Q)$, then $D(K(P),K(\delta))>D(K(Q),K(\delta))$.

Since the partition $K(P)=\{[x_i]_P|[x_i]_P=U, x_i\in U\}$ and $K(P)\prec K(Q)$, so $[x_i]_P\subseteq[x_i]_Q\subseteq U$, $x_i\in U$, and there exists $x_0\in U$ such that $[x_0]_P\subset[x_0]_Q$. Hence,

$$\begin{aligned}D(K(P),K(\delta)) &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_P\oplus U|}{|U|}=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|U|-|[x_i]_P|}{|U|}\\ &> \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|U|-|[x_i]_Q|}{|U|}=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|[x_i]_Q\oplus U|}{|U|}\\ &= D(K(Q),K(\delta)),\end{aligned}$$

That is $D(K(P),K(\delta))>D(K(Q),K(\delta))$.

Summarizing the above, $D(K(P),K(\delta))$ is an information measure. □

4.3. Partition distance and information granularity in information tables

As we know, information granularity, in a broad sense, is the average measure of information granules of a partition in a given knowledge base [43]. It can be used to characterize the classification ability of a given partition. Liang and Qian [14] developed an axiomatic definition of information granularity in information tables, which is defined as follows.

Definition 3. For any given information table $S=(U,A)$, let G be a mapping from the power set of A to the set of real numbers. We say that G is an information granularity in $S=(U,A)$ if G satisfies the following conditions:

(1) $G(P)\geqslant 0$ for any $P\subseteq A$;
(2) $G(P)=G(Q)$ for any $P,Q\subseteq A$ if there is a bijective mapping f: $K(P)\to K(Q)$ such that $|[x_i]_P|=|f([x_i]_P)|$, $x_i\in U$; and
(3) $G(P)<G(Q)$ for any P, $Q\subseteq A$ with $K(P)\prec K(Q)$.

Theorem 7. *$D(K(P),K(\omega))$ is an information granularity measure.*

Proof. Let $K(\omega)=\{[x_i]_\omega|[x_i]_\omega=\{x_i\}, x_i\in U\}$ and $K(P)=\{[x_i]_P|x_i\in U\}$.

(1) This distance D is clearly non-negative.
(2) If $K(P)=K(Q)$, then $D(K(P),K(\omega))=D(K(Q),K(\omega))$.
(3) We prove that if $K(P)\prec K(Q)$, then $D(K(P),K(\omega))<D(K(Q),K(\omega))$. From the partition $K(\omega)=\{\{x_i\}|x_i\in U\}$ and $K(P)\prec K(Q)$, one has that $\{x_i\}\subseteq S_P(x_i)\subseteq S_Q(x_i)$, $x_i\in U$, and there exists $x_0\in U$ such that $[x_0]_P\subset[x_0]_Q$. Therefore,

$$\begin{aligned}D(K(P),K(\omega)) &= \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(x_i)\oplus\{x_i\}|}{|U|}=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_P(x_i)|-1}{|U|}\\ &< \frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_Q(x_i)|-1}{|U|}=\frac{1}{|U|}\sum_{i=1}^{|U|}\frac{|S_Q(x_i)\oplus\{x_i\}|}{|U|}\\ &= D(K(Q),K(\omega)),\end{aligned}$$

i.e., $D(K(P),K(\omega))<D(K(Q),K(\omega))$.

Summarizing the above, $D(K(P),K(\omega))$ is an information granularity measure. □

Remark 1. From the above analysis, we can briefly understand information measure and information granularity by the partition distance. The bigger the value of partition distance between a partition and the finest partition is, and the higher the information content of this partition is; the smaller the value of partition distance between a partition and the coarsest partition is, and the bigger the information granularity of this partition is. In other words, the partition distance establishes the relationship between information measure and information granularity, which provides a more comprehensible perspective for uncertainty of an information table.

4.4. Relative discussions

From the above subsections, it can be seen that the partition distance can be used to information measure, which establishes a significant bridge between the partition distance and information entropy in the context of information tables. In this subsection, we will discuss the relationship between the partition distance and the heuristic functions based on information entropy for attribute reduction in rough set theory.

In order to obtain all attribute reducts of a given data set, Skowron [30] proposed a discernibility matrix method, in which any two objects determine one feature subset that can distinguish them. According to the discernibility matrix viewpoint, Qian et al. [24,25] and Shao et al. [29] provided a technique of attribute reduction for interval ordered information tables, set-valued ordered information tables and incomplete ordered information systems, respectively. The above attribute reduction methods are usually time consuming and intolerable to process large-scale data. To support efficient attribute reduction, many heuristic attribute reduction methods have been developed in rough set theory, cf. [8–10,15,16,26,31–35,38]. Slezak [32,33] investigated the relationships between information entropy, attribute clustering and attribute reduction. For convenience, from the viewpoint of heuristic functions, we classify these attribute reduction methods into four categories: positive-region reduction, Shannon's entropy reduction, Liang's entropy reduction and combination entropy reduction.

(1) Positive-region reduction
Hu and Cercone [8] proposed a heuristic attribute reduction method, called positive-region reduction, which keeps the positive region of target decision unchanged. The literature [9] gave an extension of this positive-region reduction for hybrid attribute reduction in the framework of fuzzy rough set. Jensen and Shen [11,12] proposed other extensions of the positive-region reduction to obtain an attribute reduct in the context of fuzzy rough set theory.
(2) Shannon's entropy reduction
As Shannon's information entropy was introduced to search reducts in the classical rough set model [31], Wang et al. [34] used its conditional entropy to calculate the relative attribute reduction of a decision information table. In fact, several authors also have used variants of Shannon's entropy or mutual information to measure uncertainty in rough set theory and construct heuristic algorithm of attribute reduction in rough set theory [10,35,38]. Each of these reduction methods keeps the conditional entropy of target decision unchanged.
(3) Liang's entropy reduction

Please cite this article in press as: J. Liang et al., Distance: A more comprehensible perspective for measures in rough set theory, Knowl. Based Syst. (2011), doi:10.1016/j.knosys.2011.11.003

Liang et al. [15] defined a new information entropy to measure the uncertainty of an information table and applied the entropy to reduce redundant features [16]. Unlike Shannon's entropy, this information entropy can measure both the uncertainty of an information table and the fuzziness of a rough decision in rough set theory. This reduction method can preserve the conditional entropy of a given decision table. In fact, the mutual information form of Liang's entropy also can be used to construct a heuristic function of an attribute reduction algorithm.

(4) Combination entropy reduction
In general, the objects in an equivalence class cannot be distinguished each other, but the objects in different equivalence classes can be distinguished each other in rough set theory. Therefore, in a broad sense, the knowledge content of a given attribute set can be characterized by the entire number of pairs of the objects which can be distinguished each other on the universe. Based on this consideration, Qian and Liang [26] presented the concept of combination entropy for measuring the uncertainty of information tables and used its conditional entropy to select a feature subset. This reduction method can obtain an attribute subset that possesses the same number of pairs of the elements which can be distinguished each other as the original decision table. This measure focuses on a completely different point of view, which is mainly based on the intuitionistic knowledge content nature of information gain.

The above four attribute reduction methods provide four kinds of feature subset selection approaches based on rough set theory. Positive-region reduction is to keep the positive region of target decision unchanged. From the viewpoint of partition distance, this reduction method is not based on the idea of partition distance. The other three kind of attribute reduction methods are all based on so-called conditional entropy, which can be included in information theory field. In a broad sense, owing to the conditional entropy of condition attributes with respect to a given decision attribute can be understood as the difference between the partition induced by condition attributes and that induced by the decision attribute. From this consideration, the partition distance between the condition partition and the decision partition also can play the same role. Hence, it may imply that there are some essential relationships between the partition distance and each of those heuristic functions based on conditional entropy. Owing to the importance and complexity of this task, we will carefully investigate the problem in our further study.

5. Conclusion

The contribution of this paper has two facets. On one side, through introducing a set distance to rough set theory, we have investigated how to understand measures from rough set theory in the viewpoint of distance, which are inclusion degree, accuracy measure, rough measure, approximation quality, fuzziness measure, three decision evaluation criteria, information measure and information granularity. Moreover, a rough set framework based on the set distance is also a very interesting perspective for understanding rough set approximation. On the other side, we have developed the concept of partition distance for calculating the difference between two partitions, and have used this partition distance to reveal the physical meanings of information entropy and information granularity. From the view of distance, these results look forward to providing a more comprehensible perspective for measures in rough set theory.

6. Uncited reference

[44].

Acknowledgments

The authors thank the anonymous reviewers for their constructive comments on this study.

This work was supported by the National Natural Science Foundation of China (Nos. 71031006, 60773133, 60903110 and 70971080), the National Key Basic Research and Development Program of China (973) (No. 2007CB311002), ~~and~~ the Natural Science Foundation of Shanxi Province, China (Nos. 2008011038 and 2009021017-1).

References

[1] L. An, L. Tong, Rough approximations based on intersection of indiscernibility, similarity and outranking relations, Knowledge-Based Systems 23 (6) 555–562.
[2] J. Bazan, J.F. Peters, A. Skowron, HS. Nguyen, M. Szczuka, Rough set approach to pattern extraction from classifiers, Electronic Notes in Theoretical Computer Science 82 (4) (2003) 1–10.
[3] M. Beynon, Reducts within the variable precision rough sets model: a further investigation, European Journal of Operational Research 134 (3) (2001) 592–605.
[4] C. Cornelis, M. De Cock, A. Radzikowska, Vaguely quantified rough sets, Lecture Notes in Artificial Intelligence 4482 (2007) 87–94.
[5] I. Düntsch, G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence 106 (1) (1998) 109–137.
[6] S. Greco, Z. Pawlak, R. Slowinski, Can bayesian confirmation measures be useful for rough set decision rules, Engineering Applications of Artificial Intelligence 17 (2004) 345–361.
[7] V.N. Huynh, Y. Nakamori, A roughness measure for fuzzy sets, Information Sciences 173 (2005) 255–275.
[8] X.H. Hu, N. Cercone, Learning in relational databases: a rough set approach, International Journal of Computational Intelligence 11 (2) (1995) 323–338.
[9] Q.H. Hu, Z.X. Xie, D.R. Yu, Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation, Pattern Recognition 40 (2007) 3509–3521.
[10] Q.H. Hu, D.R. Yu, Z.X. Xie, Information-preserving hybrid data reduction based on fuzzy-rough techniques, Pattern Recognition Letters 27 (5) (2006) 414–423.
[11] R. Jensen, Q. Shen, Fuzzy-rough attribute reduction with application to web categorization, Fuzzy Sets and Systems 141 (3) (2004) 469–485.
[12] R. Jensen, Q. Shen, Fuzzy-rough sets assisted attribute selection, IEEE Transactions on Fuzzy Systems 15 (1) (2007) 73–89.
[13] J.Y. Liang, D.Y. Li, Uncertainty and Knowledge Acquisition in Information Systems, Science Press, Beijing, 2005.
[14] J.Y. Liang, Y.H. Qian, Information granules and entropy theory in information systems, Sciences in China: Series F 51 (10) (2008) 1427–1444.
[15] J.Y. Liang, K.S. Chin, C.Y. Dang, C.M. Yam Richid, A new method for measuring uncertainty and fuzziness in rough set theory, International of Journal General Systems 31 (4) (2002) 331–342.
[16] J.Y. Liang, Z.B. Xu, The algorithm on knowledge reduction in incomplete information systems, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10 (1) (2002) 95–103.
[17] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publisher, London, 1991.
[18] W. Pedrycz, Relational and directional in the construction of information granules, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 32 (5) (2002) 605–614.
[19] Y.H. Qian, J.Y. Liang, Witold Pedrycz, C.Y. Dang, Positive approximation: an accelerator for attribute reduction in rough set theory, Artificial Intelligence 174 (2010) 597–618.
[20] 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.
[21] Y.H. Qian, J.Y. Liang, C.Y. Dang, Consistency measure, inclusion degree and fuzzy measure in decision tables, Fuzzy Sets and Systems 159 (2008) 2353–2377.
[22] Y.H. Qian, J.Y. Liang, C.Y. Dang, Knowledge structure, knowledge granulation and knowledge distance in a knowledge base, International Journal of Approximate Reasoning 50 (1) (2009) 174–188.
[23] G. Liu, Rough set theory based on two universal sets and its applications, Knowledge-Based Systems 23 (5) 427–433.
[24] Y.H. Qian, J.Y. Liang, C.Y. Dang, Interval ordered information systems, Computer and Mathematics with Applications 56 (2008) 1994–2009.
[25] Y.H. Qian, J.Y. Liang, C.Y. Dang, D.W. Tang, Set-valued ordered information systems, Information Sciences 179 (2009) 2809–2832.
[26] Y.H. Qian, J.Y. Liang, D.Y. Li, F. Wang, N.N. Ma, Approximation reduction in inconsistent incomplete decision tables, Knowledge-Based Systems 23 (2010) 427–433.

[27] G.F. Qiu, H.Z. Li, L.D. Xu, W.X. Zhang, A knowledge processing method for intelligent systems based on inclusion degree, Expert Systems 20 (4) (2003) 187–195.
[28] C.E. Shannon, The mathematical theory of communication, The Bell System Technical Journal 27 (3,4) (1948) 373–423.
[29] M.W. Shao, W.X. Zhang, Dominance relation and rules in an incomplete ordered information system, International Journal of Intelligent Systems 20 (2005) 13–27.
[30] A. Skowron, Extracting laws from decision tables: a rough set approach, Computational Intelligence 11 (1995) 371–388.
[31] D. Slezak, Approximate entropy reducts, Fundamenta Informaticae 53 (3–4) (2002) 365–390.
[32] D. Slezak, Rough sets and few-objects-many-attributes problem: the case study of analysis of gene expression data sets, in: Frontiers in the Convergence of Bioscience and Information Technologies, 2007, pp. 437–442.
[33] D. Slezak, Degrees of conditional (in) dependence: a framework for approximate Bayesian networks and examples related to the rough set-based feature selection, Information Sciences 179 (2009) 197–209.
[34] G.Y. Wang, H. Yu, D.C. Yang, Decision table reduction based on conditional information entropy, Chinese Journal of Computer 25 (7) (2002) 759–766.
[35] G.Y. Wang, J. Zhao, J.J. An, A comparative study of algebra viewpoint and information viewpoint in attribute reduction, Fundamenta Informaticae 68 (3) (2005) 289–301.
[36] S.K.M. Wong, W. Ziarko, Comparison of the probabilistic approximate classification and the fuzzy set model, Fuzzy Sets and Systems 21 (1986) 357–362.
[37] S.K.M. Wong, W. Ziarko, On learning and evaluation of decision rules in the context of rough sets, in: Proceedings of the International Symposium on Methodologies for Intelligent Systems, Knoxville, 1986, pp. 308–224.
[38] S.X. Wu, M.Q. Li, W.T. Huang, S.F. Liu, An improved heuristic algorithm of attribute reduction in rough set, Journal of System Sciences and Information 2 (3) (2004) 557–562.
[39] W.Z. Wu, J.S. Mi, W.X. Zhang, Generalized fuzzy rough sets, Information Sciences 151 (2003) 263–282.
[40] Z.B. Xu, J.Y. Liang, C.Y. Dang, K.S. Chin, Inclusion degree: a perspective on measures for rough set data analysis, Information Sciences 141 (2002) 227–236.
[41] W.H. Xu, W.X. Zhang, Measuring roughness of generalized rough sets induced by a covering, Fuzzy Sets and Systems 158 (2007) 2443–2455.
[42] R. Yan, J. Zhang, J. Liu, Y. Zhai, Research on the model of rough set over dual-univers, Knowledge-Based Systems 23 (8) 817–822.
[43] L.A. 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.
[44] W.X. Zhang, W.Z. Wu, J.Y. Liang, D.Y. Li, Theory and Method of Rough Sets, Science Press, Beijing, 2001.
[45] L. Zhang, B. Zhang, Fuzzy reasoning model under quotient space structure, Information Sciences 173 (2005) 353–364.
[46] W. Ziarko, Variable precision rough set model, Journal of Computer and System Science 46 (1993) 39–59.