International Journal of Computational Cognition (http://www.YangSky.com/yangijcc.htm)
Volume 1, Number 2, Pages 67–78, June 2003
Publisher Item Identifier S 1542-5908(03)10204-7/$20.00
Article electronically published on November 18, 2002 at http://www.YangSky.com/ijcc12.htm. Please cite this paper as: ⟨Jiye Liang, Zhongzhi Shi and Deyu Li, "Applications of Inclusion Degree in Rough Set Theory", *International Journal of Computational Cognition* (http://www.YangSky.com/yangijcc.htm), Volume 1, Number 2, Pages 67–78, June 2003⟩.

APPLICATIONS OF INCLUSION DEGREE IN ROUGH SET THEORY

JIYE LIANG, ZHONGZHI SHI AND DEYU LI

Abstract. Rough set theory is a relatively new mathematical tool for use in computer applications in circumstances which are characterized by vagueness and uncertainty. In this paper, applications of inclusion degree in rough set theory are discussed, the relationships among inclusion degree, measures on rough set theory and some generalized rough set models are established. These results will be very helpful for people to understand the essence of rough set theory, and can be regarded as the uniformly theoretical foundation of measures defined in rough set theory. Copyright ©2002 Yang's Scientific Research Institute, LLC. All rights reserved.

1. Introduction

Rough set theory, introduced by Z. Pawlak (see [1,2]), has often proved to be an excellent mathematical tool for the analysis of a vague description of objects. The adjective vague, referring to the quality of information, means inconsistency or ambiguity which follows from information granulation in a knowledge system. The rough sets philosophy is based on the assumption that with every object of the universe there is associated a certain amount of information (data, knowledge) expressed by means of some attributes used for object description. Objects having the same description are indiscernible with respect to the available information. The indiscernibility relation modelling the indiscernibility of objects thus constitutes a mathematical basis of rough set theory; it induces a partition of the universe into blocks of indiscernible objects, called elementary sets, that can

Received by the editors November 17, 2002 / final version received November 18, 2002.
Key words and phrases. Information system, rough sets, inclusion degree, data analysis, measure.
This work was supported by the national natural science foundation of China (No.60275019) and the natural science foundation of Shanxi, China.

be used to build knowledge about a real or abstract world. The use of the indiscernibility relation results in information granulation.

Any subset X of the universe may be expressed in terms of these blocks either precisely (as a union of elementary sets) or approximately only. In the latter case, the subset X may be characterized by two ordinary sets, called lower and upper approximations. A rough set is defined by means of these two approximations. The lower approximation of X is composed of all the elementary sets included in X (whose elements, therefore, certainly belong to X), while the upper approximation of X consists of all the elementary sets which have a non-empty intersection with X (whose elements, therefore, may belong to X). Obviously, the difference between the upper and lower approximations constitutes the boundary region of the rough set, whose elements can not be characterized with certainty as belonging or not to X, using the available information. The information about objects from the boundary region is, therefore, inconsistent or ambiguous. The cardinality of the boundary region states, moreover, to what extent it is possible to express X in exact terms, on the basis of the available information. For this reason, this cardinality may be used as a measure of vagueness of the information about X.

Rough set theory has many interesting applications. It is turning out to be methodologically significant to artificial intelligence and cognitive science, especially in the representation of and reasoning with vague and /or imprecise knowledge, machine learning, knowledge acquisition, decision analysis, knowledge discovery from databases, expert systems and pattern recognition (see[2, 7-11]). It seems of particular importance to decision support systems and data mining. The main advantage of rough set theory is that it does not need any preliminary or additional information about data, i.e., like probability in statistics, basic probability assignment in the Dempster-Shafer evidence theory, grade of membership, or the value of possibility in fuzzy set theory (see[3]). The theory is also proving to be of substantial importance in many areas of applications (see[2,7,8]).

In this paper, we discuss mainly applications of inclusion degree in rough set theory. The relationships between inclusion degree and each of various measures in rough set theory, such as degree of rough belonging, accuracy and quality of approximation of classification, dependency and importance of attributes, and accuracy and coverage of a decision rule, are established. The fact that the variable precision rough set model can also be redefined by using the concept of inclusion degree is pointed out in Section 9. These results will be very helpful for people to understand the essence of rough set theory, and can be regarded as the uniformly theoretical foundation of measures which are defined in rough set theory.

2. 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 (see [12-14]). 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 partially ordered relation was proposed in [15] for approximate reasoning. By a slight adjustment of this concept, we introduce a definition of inclusion degree into rough set data analysis.

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).

Definition 1. Let $(L, \preceq)$ be a partially ordered set. If, for any $a,\ b \in L$, there is a real number $D(b/a)$ with the following properties:
(1) $0 \leq D(b/a) \leq 1$,
(2) $a \preceq b$ implies $D(b/a) = 1$,
(3) $a \preceq b \preceq c$ implies $D(a/c) \leq D(a/b)$, and
(4) $a \preceq b$ implies $D(a/c) \leq D(b/c)$ for $\forall c \in L$,
then D is called an inclusion degree on $(L, \preceq)$.

In Definition 1, (1) is normalization for inclusion degree; (2) states the property of consistency between inclusion degree and standard inclusion; and (3) and (4) state the property of monotonicity of inclusion degree.

Inclusion degree is practically a measure on partially ordered relation, but it has more important applications than partially ordered relation.

Example 1. Let U be a finite set, $F = \{X|\ X \subseteq U\}$, and $\subseteq$ is a partially ordered relation on F. For $\forall X,\ Y \in F$, we define

$$D_0(Y/X) = \begin{cases} \frac{|Y \cap X|}{|X|}, & \text{if } X \neq \emptyset, \\ 1, & \text{if } X = \emptyset, \end{cases} \tag{1}$$

where $|X|$ denotes the cardinality of X.

It is easy to see that D_0 is an inclusion degree on $(F, \subseteq)$. In [12], D_0 is regarded as a particular case of rough inclusions.

Example 2. Let U be a finite set, G denote the set of all partitions on U, $X = \{X_1, X_2, \cdots, X_n\} \in G$ and $Z = \{Z_1, Z_2, \cdots, Z_m\} \in G$. A partially ordered relation $\preceq$ on G is defined as

$X \preceq Z$ if and only if, for $\forall X_i \in X,$ there exists $Z_j \in Z$ such that $X_i \subseteq Z_j$.

Let D_0 be an inclusion degree on $(F, \subseteq)$, where $F = \{X | \ X \subseteq U\}$. For $\forall X, Z \in G$, define

$$D_1(Z/X) = \bigwedge_{i=1}^{n} \bigvee_{j=1}^{m} D_0(Z_j/X_i), \tag{2}$$

then D_1 is an inclusion degree on $(G, \preceq)$.

In fact, we have

(1) From $0 \leq D_0(Z_j/X_i) \leq 1$, it follows that $0 \leq D_1(Z/X) \leq 1$.

(2) Let $X = \{X_1, X_2, \cdots, X_n\} \in G$, $Z = \{Z_1, Z_2, \cdots, Z_m\} \in G$ and $X \preceq Z$. For every $X_i \in X$ there exists a $Z_j \in Z$ such that $X_i \subseteq Z_j$, i.e., $\bigvee_{j=1}^{m} D_0(Z_j/X_i) = 1$, hence $D_1(Z/X) = 1$. (3) Let $X = \{X_1, X_2, \cdots, X_n\} \in G$, $Z = \{Z_1, Z_2, \cdots, Z_m\} \in G$, $Y = \{Y_1, Y_2, \cdots, Y_l\} \in G$ and $X \preceq Z \preceq Y$. Then, for every $X_i \in X$, there exist $Z_j \in Z$ and $Y_p \in Y$ such that $X_i \subseteq Z_j \subseteq Y_p$. Since D_0 is an inclusion degree on $(F, \subseteq)$, we have that

$$D_0(X_i/Y_p) \leq D_0(X_i/Z_j).$$

Hence

$$D_1(X/Y) \leq D_1(X/Z).$$

(4) Let $X = \{X_1, X_2, \cdots, X_n\} \in G$, $Z = \{Z_1, Z_2, \cdots, Z_m\} \in G$ and $X \preceq Z$. Let $Y = \{Y_1, Y_2, \cdots, Y_l\} \in G$. For arbitrary $X_i \in X$, $Z_j \in Z$ and $Y_p \in Y$ satisfying $X_i \subseteq Z_j$, since D_0 is an inclusion degree on $(F, \subseteq)$, we have

$$D_0(X_i/Y_p) \leq D_0(Z_j/Y_p).$$

Hence

$$D_1(X/Y) \leq D_1(Z/Y).$$

By Definition 1, D_1 is an inclusion degree on $(G, \preceq)$.

Rough inclusions and inclusion degree have some common characteristics on measure, but rough inclusions is more appropriate for reasoning about complex structures, inclusion degree is more appropriate for measure on partially ordered relations.

3. Basic Concepts of Rough Set Theory

Formally, an information system is an ordered quadruple $S = (U, A, V, f)$, where:

U is a non-empty finite set of objects;

A is a non-empty finite set of attributes;

V is the union of attribute domains, i.e., $V = \bigcup V_a$ for every $a \in A$, where V_a denotes the domain of the attribute a;

$f: U \times A \to V$ is an information function which associates an unique value of each attribute with every object belonging to U, i.e., $\forall a \in A$ and $x \in U$, $f(x,a) \in V_a$.

Each subset of attributes $P \subseteq A$ determines a binary indiscernibility relation $IND(P)$ as follows

$$IND(P) = \{(x,y) \in U \times U | \ \forall a \in P, f(x,a) = f(y,a)\}.$$

Obviously $IND(P)$ is an equivalence relation on the set U and

$$IND(P) = \bigcap_{a \in P} IND(\{a\}).$$

The relation $IND(P)$, $P \subseteq A$, constitutes a partition of U, which we will denote by $U/IND(P)$. Any element from $U/IND(P)$ will be called an equivalence class. Let $[x]_{IND(P)}$ denote the equivalence class of the relation $IND(P)$ containing the element x.

Let $P \subseteq A$ and $X \subseteq U$. Then $P-$lower and $P-$upper approximation of X is defined respectively as follows:

$$\underline{P}X = \bigcup\{E| \ E \in U/IND(P), E \subseteq X\}$$

and

$$\overline{P}X = \bigcup\{E| \ E \in U/IND(P), E \cap X \neq \emptyset\}.$$

The set $BN_P(X) = \overline{P}X - \underline{P}X$ will be called the $P-$boundary region of X. The set $NEG_(P)(X) = U - \overline{P}X$ will be called the $P-$ negative region of X. The set $\underline{P}X$ is the set of all elements of U, which can be with classified certainty as elements of X with respect to the values of attributes from P; the set $\overline{P}X$ consists of those elements of U which can be possibly defined as elements of X with respect to the values of attributes from P; and $BN_P(X)$ is the set of elements which can be classified neither in X nor in $U - X$ on the basis of the values of attributes from P. Finally, $NEG_P(X)$ is the set of elements which certainly do not belong in X with respect to the values of attributes from P.

For $P \subseteq A$ and $X \subseteq U$, the $P-$lower approximation of X, the $P-$upper approximation of X , the $P-$boundary of X, and the $P-$ negative region of X can be expressed by inclusion degree as follows:

$$\underline{P}X = \bigcup\{E| \ E \in U/IND(P), D_0(X/E) = 1\},$$

$$\overline{P}X = \bigcup\{E| \ E \in U/IND(P), D_0(X/E) > 0\},$$

$$BN_P(X) = \bigcup\{E| \ E \in U/IND(P), 0 < D_0(X/E) < 1\},$$

and

$$NEG_P(X) = \bigcup\{E| \ E \in U/IND(P), D_0(X/E) = 0\}.$$

4. Inclusion Degree and Accuracy Measure of Rough Set, Degree of Rough Belonging

Let $S = (U, A, V, f)$ be an information system, $P \subseteq A$, and $X \subseteq U$. The accuracy measure of rough set X with respect to P (see[1]) is defined as

$$\alpha_P(X) = \frac{|\underline{P}X|}{|\overline{P}X|}, \tag{3}$$

where $X \neq \emptyset$.

It is easy to show that

$$\alpha_P(X) = \frac{|\underline{P}X \cap \overline{P}X|}{|\overline{P}X|} = D_0(\underline{P}X/\overline{P}X).$$

The degree of rough belonging of $x \in X$ about X with respect to P (see[1]) is defined as

$$\mu_X^P(x) = \frac{|X \cap [x]_{IND(P)}|}{|[x]_{IND(P)}|}. \tag{4}$$

It follows obviously that

$$\mu_X^P(x) = D_0(X/[x]_{IND(P)}).$$

Hence, $\alpha_P(X)$ and $\mu_X^P(x)$ can be reduced to inclusion degree.

5. Inclusion Degree and Accuracy of Approximation of Classification, Quality of Approximation of Classification

Let $S = (U, A, V, f)$ be an information system, and $P \subseteq A$. Let $Y = \{Y_1, Y_2, \cdots, Y_n\}$ be a classification, or partition, of U. The origin of this classification is independent from attributes contained in P. Subsets $Y_i, i = 1, 2, \cdots, n$, are classes of classification Y. By $P-$lower and $P-$upper approximation of Y in S we mean sets $\underline{P}Y = \{\underline{P}Y_1, \underline{P}Y_2, \cdots, \underline{P}Y_n\}$ and $\overline{P}Y = \{\overline{P}Y_1, \overline{P}Y_2, \cdots, \overline{P}Y_n\}$, respectively. The coefficient

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

is called the accuracy of approximation of classification Y by the set of attributes P (see[1]), or in short, accuracy of classification. It expresses the possible correct decisions when the classified objects possess the set of

attributes P.
The coefficient

$$\gamma_P(Y) = \frac{\sum_{i=1}^{n} |\underline{P}Y_i|}{|U|} \tag{6}$$

is called the quality of approximation of classification Y by the set of attributes P (see[1]), or in short, quality of classification. It expresses the percentage of objects which can be correctly classified into class Y employing the set of attributes P.

Let $Y = \{Y_1, Y_2, \cdots, Y_n\}$ be a classification, or partition, of U. Let $F = \{\{F_1, F_2, \cdots, F_n\}|\ F_i \subseteq Y_i, i = 1, 2, \cdots, n\}$, $X = \{X_1, X_2, \cdots, X_n\} \in F$ and $Z = \{Z_1, Z_2, \cdots, Z_n\} \in F$.

A partially ordered relation $\preceq$ on F is defined as follows:

$$X \preceq Z \text{ if and only if } X_i \subseteq Z_i, \quad i = 1, 2, \cdots, n.$$

For $\forall X, Z \in F$, define

$$D_2(X/Z) = \frac{|(\bigcup_{i=1}^{n} X_i) \cap (\bigcup_{i=1}^{n} Z_i)|}{|\bigcup_{i=1}^{n} Z_i|}. \tag{7}$$

It can be easily shown that D_2 is an inclusion degree on $(F, \preceq)$.

Since $d_P(Y) = D_2(\underline{P}Y/\overline{P}Y)$ and $\gamma_P(Y) = D_2(\underline{P}Y/Y)$, $d_P(Y)$ and $\gamma_P(Y)$ can be reduced to inclusion degree.

6. Inclusion Degree and Measure of Dependency of Attributes

An information system $S = (U, A, V, f)$ can be seen as a decision table assuming that $A = C \cup D$ and $C \cap D = \emptyset$, where C is called the set of condition attributes, and D is called the set of decision attributes. Let $P \subseteq C$ and $Q \subseteq D$. The measure of dependency between P and Q (see[1]) is defined as

$$\gamma(P, Q) = \frac{|POS_P(Q)|}{|U|}, \tag{8}$$

where $POS_P(Q) = \bigcup\{\underline{P}Y|\ Y \in U/IND(Q)\}$.

Let G denote the set of all partitions on U, $X = \{X_1, X_2, \cdots, X_n\} \in G$ and $Z = \{Z_1, Z_2, \cdots, Z_m\} \in G$. A partially ordered relation $\preceq$ on G is defined as follows:

$$X \preceq Z \text{ if and only if, for } \forall X_i \in X, \text{ there exists } Z_j \in Z \text{ such that } X_i \subseteq Z_j.$$

For $\forall X, Z \in G$, define

$$D_3(Z/X) = \frac{|\bigcup_{Z_j \in Z} (\bigcup_{X_i \subseteq Z_j} X_i)|}{|U|}. \tag{9}$$

Then D_3 is an inclusion degree on $(G, \preceq)$ (see[18]).

Since $\gamma(P,Q) = D_3((U/IND(Q))/(U/IND(P)))$, $\gamma(P,Q)$ can be reduced to inclusion degree, i.e., the degree of including of partition $U/IND(Q)$ to partition $U/IND(P)$.

Remark. Let $P \to Q$ denote functional dependency between P and Q. Then $P \to Q$ if and only if $D_3((U/IND(Q))/(U/IND(P))) = 1$.

7. Inclusion Degree and Measure of Importance of Attributes

Let $S = (U, A, V, f)$ be a decision table with $A = C \cup D$ and $C \cap D = \emptyset$, where C is the set of condition attributes and D is the set of decision attributes.

The measure of importance of condition attributes $C' \subseteq C$ with respect to decision attributes D is defined as follows(see[1]):

$$\gamma(C,D) - \gamma(C - C', D). \tag{10}$$

In particular, when $C' = \{c\}$, $\gamma(C,D) - \gamma(C - \{c\}, D)$ is the measure of importance of attribute $c \subseteq C$ with respect to D.

Since

$$\begin{aligned} \gamma(C,D) - \gamma(C - C', D) &= D_3((U/IND(D))/(U/IND(C))) \\ &\quad - D_3((U/IND(D))/(U/IND(C - C'))), \end{aligned}$$

$\gamma(C,D) - \gamma(C - C', D)$ can be reduced to computation of inclusion degree.

Let D_* denote a inclusion degree D_1 or D_3 on $(G, \preceq)$. A new measure of importance of an attribute can be resulted in from D_* as follows.

Let $S = (U, A, V, f)$ be an information system. Given an attribute $a \in A$, the importance of attribute a in A is defined as

$$sig_{A-\{a\}}(a) = 1 - D_*((U/IND(A))/(U/IND(A - \{a\}))).$$

Remark. Let $S = (U, A, V, f)$ be an information system. An attribute a in A is a core attribute if and only if $sig_{A-\{a\}}(a) > 0$.

Let $S = (U, A, V, f)$ be a decision table with $A = C \cup D$ and $C \cap D = \emptyset$. Given an attribute $a \in C$, the importance of attribute a in C with respect to D is defined as

$$\begin{aligned} sig_{C-\{a\}}(a) &= 1 - D_*((U/IND(D))/(U/IND(C-\{a\}))) \\ &\quad /D_*((U/IND(D))/(U/IND(C))). \end{aligned}$$

Remark. Let $S = (U, A, V, f)$ be a decision table with $A = C \cup D$ and $C \cap D = \emptyset$. An attribute a in C with respect to D is a core attribute if and only if $sig_{C-\{a\}}(a) > 0$.

8. Inclusion Degree and Accuracy and Coverage of Decision Rule

Let $S = (U, A, V, f)$ be a decision table with $A = C \cup D$ and $C \cap D = \emptyset$, where C is the set of condition attributes and D is the set of decision attributes.

Let $U/IND(C) = \{X_1, X_2, \cdots, X_n\}$ and $U/IND(D) = \{Y_1, Y_2, \cdots, Y_m\}$ denote the partitions on U induced respectively by the equivalence relations $IND(C)$ and $IND(D)$. The expression $r_{ij} : Des_C(X_i) \to Des_D(Y_j)$ is called a $(C, D)-$ decision rule in S if $X_i \cap Y_j \neq \emptyset$, where $Des_C(X_i)$ and $Des_D(Y_j)$ are unique descriptions of the classes X_i and Y_j respectively $(i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, m)$. The set of decision rules $\{r_{ij}\}$ for each decision class Y_j $(j = 1, 2, \cdots, m)$ can be defined as

$$\{r_{ij}\} = \{Des_C(X_i) \to Des_D(Y_j) |\ Y_j \cap X_i \neq \emptyset, i = 1, 2, \cdots, n\}.$$

A decision rule r_{ij} is deterministic if and only if $Y_j \cap X_i = X_i$, and r_{ij} is non-deterministic otherwise.

The accuracy and the coverage of a decision rule r_{ij} (see [17]) are defined respectively as

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

It is notable that $\alpha_{X_i}(Y_j)$ measures the degree of sufficiency of the proposition $Des_C(X_i) \to Des_D(Y_j)$, while $\kappa_{X_i}(Y_j)$ measures the degree of its necessity. It can be easily shown that

$$\alpha_{X_i}(Y_j) = D_0(Y_j/X_i),\ \kappa_{X_i}(Y_j) = D_0(X_i/Y_j).$$

This means that both $\alpha_{X_i}(Y_j)$ and $\kappa_{X_i}(Y_j)$ can be reduced to inclusion degree.

9. Inclusion Degree and the Variable Precision Rough Set Model

Let X and Y be non-empty subsets of a finite universe U. The measure $c(X,Y)$ of the relative degree of misclassification of the set X with respect to set Y (see [16]) is defined as

$$c(X,Y) = \begin{cases} 1 - \frac{|Y \cap X|}{|X|}, & \text{if } |X| > 0, \\ 0, & \text{if } |X| = 0. \end{cases} \tag{12}$$

It can be easily shown that

$$c(X,Y) = 1 - D_0(Y/X) = D_0((U-Y)/X).$$

This means that $c(X,Y)$ can be reduced to inclusion degree.

Let $0 \leq \beta < 0.5$. Then $c(X,Y) \leq \beta$ if and only if $D_0(Y/X) \geq 1-\beta$. Thus, the variable precision rough set model (see [16]) can be expressed by inclusion degree as follows.

Let $X \subseteq U$ and R be an equivalence relation on U. The $\beta-$lower approximation of the set X is defined as

$$\underline{R}_\beta X = \bigcup \{E \in U/IND(R) |\ D_0(X/E) \geq 1-\beta\},$$

and the $\beta-$upper approximation of the set X is defined as

$$\overline{R}_\beta X = \bigcup \{E \in U/IND(R) |\ D_0(X/E) > \beta\}.$$

Consequently, the $\beta-$boundary region of X is given by

$$BNR_\beta X = \bigcup \{E \in U/IND(R) |\ \beta < D_0(X/E) < 1-\beta\}.$$

The $\beta-$negative region of X is defined as the complement of the $\beta-$upper approximation with respect to U, i.e.,

$$NEGR_\beta X = \bigcup \{E \in U/IND(R) |\ D_0(X/E) \leq \beta\}.$$

The generalized variable precision rough set model can also be expressed by inclusion degree below.

Let $0 \leq l < u \leq 1$. For any subset $X \subseteq U$, the $u-$ lower approximation of the set X is defined as

$$\underline{R}_u X = \bigcup \{E \in U/IND(R) |\ D_0(X/E) \geq u\},$$

and the $l-$ upper approximation of the set X is defined as

$$\overline{R}_l X = \bigcup \{E \in U/IND(R) |\ D_0(X/E) > l\}.$$

The $(l,u)-$ boundary region of X is given by

$$BNR_{l,u} X = \bigcup \{E \in U/IND(R) |\ l < D_0(X/E) < u\}.$$

The $l-$negative region of X is defined as

$$NEGR_l X = \bigcup \{E \in U/IND(R) | \ D_0(X/E) \leq l\}.$$

10. Conclusions

Rough set theory is an important mathematical tool to deal with vagueness and uncertainty in knowledge systems. In this paper, we discuss mainly applications of inclusion degree in rough set theory, establish the relationships between the inclusion degree and each of various measures in rough set theory, such as degree of rough belonging, accuracy and quality of approximation of classification, dependency and importance of attributes, and accuracy and coverage of a decision rule and so on. The variable precision rough set model and the generalized variable precision rough set model is redefined by using the concept of inclusion degree. These results will be very helpful for people to understand the essence of rough set theory, and can be regarded as the uniformly theoretical foundation of the measures which are defined in rough set theory. The concept of inclusion degree will play a significant role in further research on rough set theory.

References

[1] Z. Pawlak, Rough sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991.

[2] Z. Pawlak, J.W. Grzymala-Busse, R. Slowiński, W. Ziarko, Rough sets, Comm. ACM 38: 89-95, 1995.

[3] Z. Pawlak, Rough set theory and its application to data analysis, Cybernetics and Systems: An International Journal 29: 661-688, 1998.

[4] I. Düntsch, G. Gediga, Uncertainty measures of rough set prediction, Artificial Intelligence 106: 109-137, 1998.

[5] J.W. Guan, D.A. Bell, Rough computational methods for information systems, Artificial Intelligence 105: 77-103, 1998.

[6] D.A. Bell, J.W. Guan, Computational methods for rough classification and discovery, Journal of the American Society for Information Science 49: 403-414, 1998.

[7] L. Polkowski, A. Skowron, (Eds.), Rough Sets in Knowledge Discovery 1. Methodology and Applications, Physica-Verlag, 1998.

[8] L. Polkowski, A. Skowron, (Eds.), Rough Sets in Knowledge Discovery 2. Applications, Case Studies and Software Systems, Physica-Verlag, 1998.

[9] M. Kryszkiewicz, Rough set approach to incomplete information systems, Information Sciences 112: 39-49, 1998.

[10] J.Y. Liang, Z.B. Xu, The algorithm on knowledge reduction in incomplete information systems, International Journal of Unvertainty, Fuzziness and Knowledge-Based Systems, 24: 95-103, 2002.

[11] P.J. Lingras, Y.Y. Yao, Data mining using extensions of the rough set model, Journal of American Society for Information Science, 49: 415-422, 1998.

[12] L.Polkowski, A.Skowron, Rough mereology, Lecture Notes in Artificial Intelligence Vol.869, Springer-Verlag, Berlin, 1994. page 85-94.

[13] L. Polkowski, A. Skowron, Rough mereology: a new paradigm for approximate reasoning, International Journal of Approximate Reasoning 15: 333-365, 1996.
[14] L. Polkowski, T.Y. Lin, S. Tsumoto (Eds.), Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems, Physica-Verlag, Heidelberg, 2000.
[15] W.X. Zhang, Y. Leung, Theory of including degrees and its applications to uncertainty inferences, Soft Conputing in Intelligent Systems and Information processing, Institute of Electrial and Electronics Engineers, New York, 1996. page 496-501.
[16] W.Ziarko, Variable precision rough set model, Journal of Computer and System Sciencs 46: 39-59, 1993.
[17] S. Tsumoto, Extraction of experts' decision rule from clinical databases using rough set model, Intelligent Data Analysis 2: 215-227, 1998.
[18] 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: 227-236, 2002.

Jiye Liang and Zhongzhi Shi, Institut of Computing Technology, Chinese Academy of Science, Beijing, 100080, China.
Jiye Liang and Deyu Li, Department of Computer Science, Shanxi University , Taiyuan, 030006, China.
E-mail address: ljy@sxu.edu.cn (Jiye Liang)