Binary Relations as a Basis for Rule Induction in Presence of Quantitative Attributes

Liping An
Business School, Nankai University, Tianjin, 300071, China
Email: anliping2000@sina.com

Lingyun Tong
School of Management, Hebei University of Technology, Tianjin, 300130, China
Email: tonglingyun2008@sina.com

Abstract—In original rough set theory, the notion of set approximation has been introduced by using indiscernibility relation defined on the set of objects. In some cases, it is necessary to generalize indiscernibility relation by using some other binary relations. In this paper, we consider similarity relations and tolerance relations among objects. These binary relations are defined from some similarity measures at the level of values of any quantitative attribute. The relations defined by single attribute are aggregated into a global relation at the level of the set of attributes. Then, we construct the lower approximation operation and the upper approximation operation generated by a binary relation and its inverse relation. In order to induce the minimal decision rules used to support the decision task, the nonsimilarity matrix of a decision table with respect to the lower approximation and boundary is defined to construct the nonsimilarity functions which are Boolean functions. The set of “if … then …” decision rules is decoded from prime implicants of the Boolean functions. An example is illustrated to demonstrate the application of this approach.

Index Terms—binary relations, similarity, tolerance, rough sets, lower and upper approximations

I. INTRODUCTION

The concept of rough set was originally proposed by Pawlak [1] as a mathematical approach to handle imprecision, vagueness and uncertainty in data analysis. A rough set is defined by means of two ordinary sets called lower and upper approximations. The indiscernibility classes identified by indiscernibility relation are the building blocks for the construction of the lower and upper approximations. The requirement of an indiscernibility relation seems to be a stringent condition that may limit the application domain of the Pawlak rough set model [2]. For example, this kind of relation implies that attributes are all nominal, which have exemplified unreasonable results for human intuition when some attributes are ordinal or quantitative [3]. And, incomplete information systems [4, 5] can not be handled with Pawlak’s rough sets. To overcome the unreasonablenss, several interesting and meaningful extensions to indiscernibility relation have been proposed to solve these problems. Stepaniuk et al [8] proposed a tolerance relation, which is reflexive and symmetric, on attributes values in information system. Skowron and Stepaniuk [9] presented a more general definition of approximation space which can be used for example for similarity based rough set model and variable precision rough set model. Stepaniuk [10] further discussed rough set model based on approximation spaces with uncertainty functions and rough inclusions. The elements of approximation space are parametrized. And, the strategies of parameter optimization were discussed. Kryszkiewicz [4] defined a tolerance relation in incomplete information systems and proposed a type of attribute reduction that only eliminates the information, which is not essential from the point of view of classification or decision making. Exploring the idea of discernibility functions, decision rules are found directly from the incomplete decision table. Greco et al. [11] introduced the rough approximation based on dominance relations and proposed a rough set methodology to analyze multi-criteria choice and ranking decision problems. Slowinski and Vanderpooten [12] proposed the substitution of an equivalence relation with a general relation in which only reflexivity is required. Dubois and Prade [13] combined fuzzy sets with rough sets in a fruitful way by defining rough fuzzy sets and fuzzy rough sets. Chen et al. [14] replaced the indiscernibility relation by a fuzzy similarity relation and generalized the crisp rough sets to fuzzy rough sets. In [15, 16], Yao investigated rough sets based on general binary relations. The author started from the properties of binary relations to investigate the essential properties of the lower and upper approximation operations generated by such relations. His studies are mainly concentrated on the constructive and axiomatic approaches of approximation operators. In [17, 18], the concept of a cover of a universe was proposed to construct the upper and lower approximations of an arbitrary set. Zhu [19] studied covering-based rough sets from the topological point of view.
In reality, due to the imprecision of data describing the objects, small differences are often not considered significant for the purpose of discrimination. This situation may be formally modeled by considering similarity or tolerance relations [12, 20-23]. Extending indiscernibility to similarity or tolerance imposes to weakening of some of the properties of the binary relation in terms of symmetry and transitivity.

In this paper, the binary relations are defined from some similarity measures at the level of values of any single attribute. Then the relations defined by single attribute are aggregated into a global relation at the level of the set of attributes. Set approximation can be defined based on the global relations. In order to induce the minimal decision rules used to support the decision task, the nonsimilarity matrix of a decision table with respect to the lower approximation and boundary is defined to construct the nonsimilarity functions which are Boolean functions. The set of "if … then … " decision rules is decoded from prime implicants of the Boolean functions.

The other parts of this paper are organized as follows. Section 2 reviews some basic notions of binary relations and inverse relations. In Section 3, we discuss the binary relations on quantitative attribute values in information tables and the rough approximations based on the relations. Section 4 presents a method for decision rule induction. The rule induction algorithm is illustrated with an example in Section 5. The final Section contents conclusions.

II. BINARY RELATION AND ITS INVERSE RELATION

In this section, we present some basic concepts and properties of binary relations [12, 16, 24, 25] to be used in this paper.

Definition 1. Let U be a non-empty finite set. U×U is the product set of U and U. Any subset R of U×U is called a binary relation on U. For any (xₙ, xₗ)∈U×U, if (xₙ, xₗ)∈R, we say xₙ has relation R with xₗ and denote this relationship as R(xₙ, xₗ). The relation R is called an indiscernibility relation on U, denoted by R⁻¹, if R⁻¹={((xₙ, xₗ)): xₙRxₗ}.

Definition 2. Let R⊆U×U be a binary relation on U. The inverse relation of R, denoted by R⁻¹, is defined as R⁻¹={((xₙ, xₗ)): xₗRxₙ}.

Definition 3. For any xₙ∈U, we call the set R(xₙ)={xₗ∈U: xₗRxₙ} the left neighborhood of xₙ in R.

Definition 4. For any xₙ∈U, we call the set R⁻¹(xₙ)={xₗ∈U: xₗRxₙ} the right neighborhood of xₙ in R.

From the Definition 1–4, the following proposition can easily be obtained.

Proposition 1. xₙRxₗ↔xₗRxₙ∈R(xₙ, xₗ)xₗRxₙ∈R⁻¹(xₙ) xₗRxₙ∈R⁻¹(xₙ).

Definition 5. Let R⊆U×U be a binary relation on U. The relation R is said to be serial if there exists xₙ∈U such that (xₙ, xₙ)∈R for all xₙ∈U. R is said to be reflexive if (xₙ, xₙ)∈R, or xₙRxₚ or x₉Rx₅ for all xₙ∈U; R is said to be symmetric if for all xₙ, xₗ∈U, (xₙ, xₗ)∈R↔(xₗ, xₙ)∈R⁻¹; R is said to be transitive if for all xₙ, xₗ, xₚ∈U, (xₙ, xₗ)∈R→(xₙ, xₚ)∈R⁻¹, (xₚ, xₗ)∈R, (xₙ, xₕ)∈R→(xₕ, xₗ)∈R⁻¹, (xₕ, xₙ)∈R⁻¹, (xₕ, xₕ)∈R⁻¹.

Proof. Let xₙRxₗ and xₗRxₚ. Then xₙRxₚ.

Proposition 6. Let R be a relation on U. If R is reflexive, we say R is a binary relation on U. Proposition 7. Let R be a relation on U. If R is reflexive, we say R is a similarity relation on U.

The standard rough set model can be generalized by considering any type of binary relation on condition attribute values instead of the indiscernibility relation. In order to construct a binary relation, one can start from setting relations between attribute values for each attribute. Depending on additional information about an attribute, similarity measures can be created. Relationship between the binary relation Rₙ and similarity measure Sₙ could be described as follows [26].

xₙRxₗ↔Sₙ(xₙ, xₗ)>t(a),

where a is a value attribute.
where \( a \in A \), \( x_i, x_k \in U \). \( R_a \) is a binary relation between attribute values of attribute \( a \), \( t(a) \) is a similarity threshold for values of attribute \( a \) and \( t(a) \in [0, 1] \).

Similarity measure and distance are closely related. Using the distance measure \( d_a(x_i, x_k) \) on attribute values, two values \( f(x_i, a) \) and \( f(x_k, a) \) are considered to be similar if and only if

\[
d_a(x_i, x_k) \leq t(a),
\]

where \( t(a) \) is a threshold for values of attribute \( a \) and \( t(a) \in [0, 1] \).

Then, binary relations can be obtained on attribute values by using distance measures. And the following statement hold:

\[
x_i R_a x_k \iff d_a(x_i, x_k) \leq t(a).
\]

### B. Binary Relations on Quantitative Attribute Values

Binary relations on quantitative attribute values are constructed on the basis of that small differences on some attribute values may be judged unimportance. Tolerance relations and similarity relations are considered for the quantitative attributes.

Let \( S = \{ U, A, V, f \} \) be a decision table. \( a \in A \) is a quantitative attribute. The distance between values \( f(x_i, a) \) and \( f(x_k, a) \) can be defined as:

\[
d_a(x_i, x_k) = \frac{|f(x_i, a) - f(x_k, a)|}{\max(f(x_i, a)) - \min(f(x_i, a))}.
\]  

**Definition 8.** Let \( S = \{ U, A, V, f \} \) be an information table. \( a \in A \) is a quantitative attribute. The binary relation of \( a \) on \( U \), denoted by \( R_a^+ \), is defined as

\[
R_a^+ = \{(x_i, x_k) \in U \times U : d_a(x_i, x_k) \leq t(a)\}.
\]  

When \( (x_i, x_k) \in R_a^+ \), we say that \( x_i \) and \( x_k \) are similar on attribute \( a \), denoted by \( x_i R_a^+ x_k \). Obviously, \( R_a^+ \) is a tolerance relation.

For review of different distance measures defined on attribute values see [27].

Slowinski and Vanderpooten [12, 22] have proposed a similarity relation which is only reflexive.

**Definition 9.** Let \( S = \{ U, A, V, f \} \) be an information table. \( a \in A \) is a quantitative attribute. The binary relation of \( a \) on \( U \), denoted by \( R_a^- \), is defined as

\[
R_a^- = \{(x_i, x_k) \in U \times U : |f(x_i, a) - f(x_k, a)| \leq \varepsilon_a(f(x_i, a))\},
\]  

where \( x_i \) is a subject object and \( x_k \) is a referent object. \( \varepsilon_a \) is a function of the value of attribute \( a \) for the referent object \( x_k \). In practice, the following linear form of \( \varepsilon_a \) is sufficient to characterize insignificant differences [22]:

\[
\varepsilon_a(f(x_i, a)) = \alpha_a f(x_i, a) + \beta_a.
\]  

Except for constant thresholds corresponding to \( \alpha_a = 0 \), \( R_a^- \) defined according to (3) are not symmetric. When \( \alpha_a = \beta_a = 0 \), \( R_a^- \) is an indiscernibility relation.

According to (3) and (4), \( x_i \) is similar to \( x_k \) on attribute \( a \) if

\[
|f(x_i, a) - f(x_k, a)| \leq \varepsilon_a(f(x_i, a)),
\]

i.e.,

\[
f(x_i, a) - \varepsilon_a(f(x_i, a)) \leq f(x_k, a) \leq f(x_i, a) + \varepsilon_a(f(x_i, a)).
\]

**Definition 10.** For any \( x_i \in U \), we call the interval

\[
[f(x_i, a) - \varepsilon_a(f(x_i, a)), f(x_i, a) + \varepsilon_a(f(x_i, a))]
\]

the similarity class interval of \( x_i \) on attribute \( a \), denoted by

\[
[x_i^-(a), x_i^+(a)].
\]

Then, the following relation holds:

\[
x_i R_a^- x_k \iff f(x_i, a) \in [x_k^-(a), x_k^+(a)].
\]

When \( (x_i, x_k) \in R_a^- \), we say that \( x_i \) and \( x_k \) are similar on attribute \( a \), denoted by \( x_i R_a^- x_k \). \( R_a^- \) is a similarity relation [22].

### C. Binary Relations and Rough Approximations

The relations defined by single attribute are aggregated into a global relation at the level of the set of attributes.

**Definition 11.** Let \( S = \{ U, A, V, f \} \) be an information table. The binary relation between objects with respect to the set of attributes \( B \subseteq A \), denoted by \( R_B \), is defined as

\[
R_B = \{(x_i, x_k) : x_i R_B x_k, \forall b \in B\}.
\]

If \( (x_i, x_k) \in R_B \), it is said that the object \( x_i \) is similar to \( x_k \) with respect to \( B \), denoted by \( x_i R_B x_k \).

The following relation hold:

\[
R_B = \bigcap_{b \in B} R_b.
\]

**Definition 12.** The inverse relation of \( R_B \), denoted by \( R_B^- \), is defined as

\[
R_B^- = \{(x_i, x_k) : x_i R_B x_k, \forall b \in B\}.
\]

\( R_B(x_i) = \{x_k \in U : x_i R_B x_k\} \) is the set of subject object \( x_i \) which are similar to \( x_i \) with respect to \( B \).

\( R_B^-(x_i) = \{x_k \in U : x_i R_B x_k\} \) is the set of referent object \( x_i \) to which \( x_i \) is similar with respect to \( B \).

**Definition 13.** Let \( S = \{ U, A, V, f \} \) be an information table. \( X \subseteq U \) and \( \emptyset \neq B \subseteq A \). \( R_B \) is a binary relation defined on \( U \). The \( B \)-lower approximation of \( X \) in \( S \) with respect to \( R_B \), denoted by \( B(X) \), and the \( B \)-upper approximation of \( X \) in \( S \) with respect to \( R_B \), denoted by \( \overline{B}(X) \), are respectively defined as follows:

\[
B(X) = \{x \in U : R_B(x) \cap X \neq \emptyset\},
\]

\[
\overline{B}(X) = \{x \in U : R_B(x) \cap X \neq \emptyset\}.
\]

The \( B \)-boundary of \( X \) in \( S \) with respect to \( R_B \) is \( B(X) - \overline{B}(X) \), denoted by \( Bn_B(X) \).
Definition 14. The $B$-lower approximation of $X$ in $S$ with respect to $R^+_{a}$, denoted by $\overline{B}^+(X)$, and the $B$-upper approximation of $X$ in $S$ with respect to $R^+_a$, denoted by $\overline{B}^{-1}(X)$, are respectively defined as follows:

$$\overline{B}^+(X) = \{x \in U: R^+_a(x) \subseteq X\}, \quad (10)$$

$$\overline{B}^{-1}(X) = \{x \in U: R^+_a(x) \cap X \neq \emptyset\}. \quad (11)$$

The $B^+$-boundary of $X$ in $S$ with respect to $R^+_a$ is $\overline{B}^{-1}(X) - \overline{B}^{-1}(X)$, denoted by $B^{-1}(X)$.

If $R_B$ is a binary relation which is symmetric, such as a tolerance relation, then $B(X) = \overline{B}^{-1}(X)$, $\overline{B}(X) = \overline{B}^{-1}(X)$.

IV. DECISION RULE INDUCTION

Decision rules are logical statements of the type “if...then...”, where the antecedent specifies values assumed by one or more condition attributes and the consequent specifies an assignment to one or more decision classes. Exact rules are supported only by the lower approximation of the corresponding decision class. Approximate rules are supported only by objects from the boundaries of the corresponding decision classes. Procedures for generation of decision rules from a decision table use an inductive learning principle. The objects are considered as examples of decisions. In order to induce decision rules with a unique consequent assignment to an elementary set, the examples belonging to the elementary set are called positive and all the others negative. A decision rule is discriminant if it distinguishes positive examples from negative ones, and minimal, i.e. removing any attribute from a condition part gives a rule covering also negative objects [11].

A. Nonsimilarity Matrix and Functions

In classical rough set theory, all minimal decision rules can be generated using indiscernibility matrix and its Boolean functions [28]. For the rough approximations by means of similarity relation, the concepts of nonsimilarity matrix and nonsimilarity functions are presented in this paper to induce all the minimal decision rules.

Definition 15. Let $S=(U, C \cup \{d\}, V, f)$ be a decision table, $\emptyset \neq B \subseteq C$. $U_q$ is the $q$-th decision class. The nonsimilarity matrix $M_d(B_n(U_q))$ of $S$ with respect to the $B$-boundary of $U_q$ is defined as

$$M_d(B_n(U_q))=\{m(x, y)\}_{i,j}, \quad (12)$$

where $i=1, 2, \ldots, I$, $j=1, 2, \ldots, J$.

$$m(x_i, y_j)=\{b_n \in B: \text{not } x_i R_n y_j \text{ and } x_i \in B(U_q) \cap \overline{B}^{-1}(U_q),$$

$$y_j \in U - (B(U_q) \cap \overline{B}^{-1}(U_q)).$$

Definition 16. The nonsimilarity function $f_{B_n(U_q)}(x_i)$ of $x_i$ from $M_d(B(U_q) \cap \overline{B}^{-1}(U_q))$ is defined by:

$$f_{B_n(U_q)}(x_i) = \bigwedge_{j_y} b_i^*: b_n \in m(x, y) \text{ and } m(x, y) \neq \emptyset, \quad (13)$$

where $b_i^*$ is the Boolean variable corresponding to the attribute $b_n$. $\land$ and $\lor$ are respectively generalized conjunction and disjunction operators.

Turning $f_{B_n(U_q)}(x_i)$ into disjunctive normal form and using the absorption law of Boolean algebra to simplify it, the conjuncts, i.e., the prime implicants of the simplified decision function correspond to the minimal exact decision rules.

Definition 17. Let $S=(U, C \cup \{d\}, V, f)$ be a decision table, $\emptyset \neq B \subseteq C$. $U_q$ is the $q$-th decision class. The nonsimilarity matrix $M_d(B_n(U_q))$ of $S$ with respect to the $B$-boundary of $U_q$ is defined as

$$M_d(B_n(U_q))=\{m(x_i, y_j)\}_{i,j}, \quad (14)$$

where $i=1, 2, \ldots, I$, $j=1, 2, \ldots, J$.

$$m(x_i, y_j)=\{b_n \in B: \text{not } x_i R_n y_j \text{ and } x_i \in B_n(U_q),$$

$$y_j \in U - B_n(U_q).$$

Definition 18. The nonsimilarity function $f_{B_n(U_q)}(x_i)$ of $x_i$ from $M_d(B_n(U_q))$ is defined as

$$f_{B_n(U_q)}(x_i) = \bigwedge_{j_y} b_i^*: b_n \in m(x, y) \text{ and } m(x, y) \neq \emptyset, \quad (15)$$

Turning $f_{B_n(U_q)}(x_i)$ into disjunctive normal form and using the absorption law of Boolean algebra to simplify it, the prime implicants of the simplified decision function correspond to the minimal approximation decision rules.

When we turn the nonsimilarity function (13) or (15) into disjunctive normal form, the following types of rules can be obtained:

$$\bigwedge_{a} (\text{not } x_i R_n y_j) \rightarrow x_i \in \cup_{q} U_q, \quad \text{i.e.,} \quad \bigwedge_{a} (f(x_i, b_n) \in \{ y_j^a (b_n) \}) \rightarrow x_i \in \cup_{q} U_q. \quad (16)$$

Since $x_i R_n y_j$, we have

$$f(x_i, b_n) \in \{ x_i^a (b_n) \}. \quad (17)$$

By (16) and (17), we get the following statement.

$$\bigwedge_{a} (f(x_i, b_n) \in \{ y_j^a (b_n) \}) \land (f(x_i, b_n) \in \{ x_i^a (b_n) \}) \rightarrow x_i \in \cup_{q} U_q. \quad (18)$$

Similarly, we can define the $M_d(B(U_q))$ of $S$ with respect to the $B^{-1}$-boundary of $U_q$ and $f_{B_n(U_q)}(x_i)$.

B. Algorithm for Rule Induction

The algorithm of rule induction from a decision table presented is in the following way.

Input: A decision table $S=(U, C \cup \{d\}, V, f)$.

Output: Set of all minimal rules that distinguish objects belonging to $C(U_q) \cap \overline{C}^{-1}(U_q)$, $B_n(U_q)$ and $B^{-1}(U_q)$.
Algorithm:

Step 1. For each qualitative attribute \( b_i \in C \), compute the indiscernibility class \( I_{b_i}(x_i) \) of each object \( x_i \) in the decision table, i.e.,

\[ I_{b_i}(x_i) = \{ x_i \in U : x_iI_{b_i}x_i \} \]

Step 2. For each quantitative attribute \( e_i \in C \), compute the similarity class intervals of each object \( x_i \) in the decision table according to (5).

Step 3. For each object \( x_i \), compute the similarity class for each quantitative attribute \( e_i \):

\[ R_{e_i}(x_i) = \{ x_i \in U : x_iR_{e_i}x_i \} \]

Step 4. For all of the qualitative and quantitative attributes, aggregate the indiscernibility classes \( I_{b_i}(x_i) \) and the similarity classes \( R_{e_i}(x_i) \) of object \( x_i \) using the intersection operator, i.e.,

\[ R_C(x_i) = \bigcap_{b_i \in C} I_{b_i}(x_i) \bigcap_{e_i \in C} R_{e_i}(x_i) \]

Step 5. Compute the set of objects to which \( x_i \) is similar according to \( R_C \), i.e.,

\[ R_C^{-1}(x_i) = \{ x_i \in U : x_iR_Cx_i \} \]

Step 6. For each decision class \( U_{C_i} \), compute the \( C(U_{C_i}) \) and \( Bn_C(U_{C_i}) \) and \( C^{-1}(U_{C_i}) \) and \( Bn_C^{-1}(U_{C_i}) \).

Step 7. Construct the nonsimilarity matrix \( M_S(C(U_{C_i}) \cap C^{-1}(U_{C_i})) \) of \( S \) with respect to the upper approximation of \( U_{C_i} \).

Step 8. For each object \( x_i \in U_{C_i} \), \( i = 1, 2, \ldots, I \), construct the nonsimilarity function \( f_{C(U_{C_i})}(x_i) \) for \( x_i \) from \( M_S(C(U_{C_i}) \cap C^{-1}(U_{C_i})) \).

Step 9. Construct the nonsimilarity matrix \( M_S(Bn_C(U_{C_i})) \) of \( S \) with respect to the boundary of \( U_{C_i} \).

Step 10. For each object \( x_i \in U_{C_i} \), \( i = 1, 2, \ldots, I \), construct the nonsimilarity function \( f_{Bn_C(U_{C_i})}(x_i) \) of \( x_i \) from \( M_S(Bn_C(U_{C_i})) \).

Step 11. Similarly, construct the nonsimilarity matrix \( M_S(Bn_C^{-1}(U_{C_i})) \) of \( S \) with respect to the boundary of \( U_{C_i} \) and \( f_{Bn_C^{-1}(U_{C_i})}(x_i) \).

Step 12. Calculate the disjunctive normal form of \( f_{C(U_{C_i})}(x_i) \), \( f_{Bn_C(U_{C_i})}(x_i) \) and \( f_{Bn_C^{-1}(U_{C_i})}(x_i) \) where each conjunct corresponds to a minimal decision rule defined as (18).

Step 13. Repeat the Step 6 to 12 to obtain the decision rules of all the objects and simplify the rule set.

V. AN EXAMPLE

In order to illustrate the rule generation algorithm, let us consider a simple example of a hypothetical decision table, as shown in Table I. The objects are characterized by two condition attributes \( a \) and \( b \). \( a \) is a qualitative attribute, while attribute \( b \) is a qualitative one with three possible values. Let \( C = \{ a, b \} \). Decision attribute \( d \) makes a dichotomic partition of the objects. The decision classes \( U_1 \) and \( U_2 \) are, respectively:

\[ U_1 = \{ x \in U : d(x) = 1 \} = \{ x_1, x_3, x_5, x_8, x_{11} \} \]
\[ U_2 = \{ x \in U : d(x) = 2 \} = \{ x_2, x_4, x_6, x_9, x_{10}, x_{12} \} \]

Let us use the similarity relation defined according to (3) and (4) for the condition attribute \( a \). Suppose that \( \alpha = 0.2, \beta = 0 \). The similarity relation on attribute \( a \) is defined as

\[ R_a = \{ (x_i, x_i') \in U \times U : |f(x_i, a) - f(x_i', a)| \leq 0.2 \} \]

Suppose that the binary relation on attribute \( b \) is an indiscernibility relation:

\[ I_b = \{ (x_i, x_i') \in U \times U : f(x_i, b) = f(x_i', b) \} \]

For attribute \( a \), the similarity class interval of \( x_i \) are computed as follows.

\[ \{ x_i^+(a), x_i^-(a) \} = [43 - (0.2 \times 43 + 0), 43 + (0.2 \times 43 + 0)] = [34.4, 51.6] \]

The similarity class intervals of \( x_i \), \( i = 1, 2, \ldots, 12 \), on attribute \( a \) are shown in the second column of Table II. Then, the similarity class of \( x_i \) for attribute \( a \) is shown in the third column. The indiscernibility classes of \( x_i \) for qualitative attribute \( b \) are shown in the fourth column. The similarity class \( R_C(x_i) \) of \( x_i \) for attributes \( a \) and \( b \) can be aggregated using intersection operator, i.e.,

\[ R_C(x_i) = R_a(x_i) \cap I_b(x_i) \]

as shown in the fifth column. The class of objects to which \( x_i \) is similar, i.e., \( R_C^{-1}(x_i) \), are shown in the sixth column of Table II.

Then, the approximations of decision classes and the boundary region of \( U_1 \) and \( U_2 \) are, respectively:

\[ C(U_1) = \{ x_1, x_3, x_5 \} \]
\[ \overline{C}(U_1) = \{ x_2, x_4, x_6, x_9, x_{11}, x_{12} \} \]
\[ Bn_C(U_1) = \{ x_2, x_4, x_6, x_{11}, x_{12} \} \]
\[ \overline{C}(U_2) = \{ x_2, x_4, x_6, x_{11}, x_{12} \} \]

### Table I. A Decision Table

<table>
<thead>
<tr>
<th>Firm</th>
<th>a</th>
<th>b</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>x_1</td>
<td>43</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>x_2</td>
<td>54</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>x_3</td>
<td>124</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>x_4</td>
<td>102</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>x_5</td>
<td>98</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>x_6</td>
<td>88</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>x_7</td>
<td>130</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>x_8</td>
<td>128</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>x_9</td>
<td>82</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>x_{10}</td>
<td>134</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>x_{11}</td>
<td>58</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>x_{12}</td>
<td>126</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
corresponding minimal decision rule is the following:

\[(\text{example}, (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12})) \text{ are constructed to obtain the minimal rules. For} \]

\[
\begin{align*}
\text{The nonsimilarity matrix } M_d(\overline{C}(U_2) \cap C^{-1}(U_1)) & \text{ is} \\
\text{constructed to obtain the} \text{ minimal rules. For} \end{align*}
\]

\[
\begin{align*}
\text{By } M_d(\overline{C}(U_1) \cap C^{-1}(U_1)), \text{ the nonsimilarity functions } f_{\overline{C}(U_2)}(x_i) \text{ are constructed to obtain the minimal rules. For example,} \\
f_{\overline{C}(U_1)}(x_i) = a \land (a \land b) \land (a \land b) \land (a \land b) \land (a \land b) \land (a \land b).
\end{align*}
\]

\[
\text{Turn } f_{\overline{C}(U_1)}(x_i) \text{ into disjunctive normal form, we have} \\
f_{\overline{C}(U_1)}(x_i) = a,
\]

i.e., the prime implicant of \( f_{\overline{C}(U_1)}(x_1) \) is \( a \).

The corresponding minimal decision rule is the following:

\[
\begin{align*}
f(x_1, a) \in [34.4, 43.2) \rightarrow f(x_1, d) = 1. \\
\text{Similarly,} \\
f(x_1, a) \in (122.4, 156] \rightarrow f(x_1, d) = 1.
\end{align*}
\]

The nonsimilarity matrix \( M_d(\overline{C}(U_2) \cap C^{-1}(U_1)) \) is computed as follows:

\[
\begin{align*}
\text{Since } f_{\overline{C}(U_1)}(x_i) = b, \text{ where } i = 5, 6, 10, \text{ the following rule can be obtained:} \\
f(x_1, b) = 2 \rightarrow f(x_1, d) = 2.
\end{align*}
\]

Therefore, three exact rules supported by objects from the lower approximation of the corresponding decision class are as follows:

\[
\begin{align*}
f(x_1, a) \in [34.4, 43.2) \rightarrow f(x_1, d) = 1. (x_1) \\
f(x_1, a) \in (122.4, 156] \rightarrow f(x_1, d) = 1. (x_7) \\
f(x_1, b) = 2 \rightarrow f(x_1, d) = 2. (x_5, x_6, x_{10})
\end{align*}
\]

Similarly, \( M_d(B_n(U_2)) \) and \( M_d(B_n^{-1}(U_2)) \), \( q=1, 2, \) are computed as follows, respectively:

\[
\begin{align*}
x_1 & \quad x_2 & \quad x_3 & \quad x_4 & \quad x_5 & \quad x_6 & \quad x_7 & \quad x_8 & \quad x_9 & \quad x_{10} & \quad x_{11} & \quad x_{12} \\
\text{15, 6, 9, 6} & \quad \text{ab} & \quad \text{ab} & \quad \text{ab} & \quad \text{ab} & \quad \text{ab} & \quad \text{ab} & \quad \text{ab} & \quad \text{ab} & \quad \text{ab} & \quad \text{ab} & \quad \text{ab}
\end{align*}
\]


Since $Bn_C(U_1) = \{x_2, x_3, x_4, x_6, x_{11}, x_{12}\}$ and $Bn_C(U_2) = \{x_2, x_3, x_4, x_5, x_9, x_{11}, x_{12}\}$, we have

$$M_\delta(Bn_C(U_2)) = M_\delta(Bn_C(U_1))$$

and

$$M_\delta(Bn_C(U_3)) = M_\delta(Bn_C(U_1)).$$

Then, the approximate rules supported by objects from the boundaries of the corresponding decision classes are as follows:

$$f(x, a) \in (51.6, 64.8) \Rightarrow \neg f(x, d) = \neg \neg f(x, d) = 2, (x_3)$$

$$f(x, a) \in (117.6, 148.8) \wedge f(x, b) = 1 \Rightarrow \neg f(x, d) = \neg \neg f(x, d) = 2, (x_3)$$

$$f(x, b) = 1 \Rightarrow \neg f(x, d) = 1 \wedge f(x, d) = 2, (x_3, x_4, x_5, x_6, x_9, x_{12})$$

$$f(x, a) \in (51.6, 64.8) \Rightarrow \neg f(x, d) = 1 \wedge f(x, d) = 2, (x_1)$$

Some of the rules, induced from the different objects, may be simplified, e.g.,

$$f(x, a) \in (51.6, 64.8) \Rightarrow f(x, d) = 1 \wedge f(x, d) = 2$$

and

$$f(x, a) \in (51.6, 65.6) \Rightarrow f(x, d) = 1 \wedge f(x, d) = 2.$$  

We can unite them into one rule of

$$f(x, a) \in (51.6, 64.8) \Rightarrow f(x, d) = 1 \wedge f(x, d) = 2.$$  

The final rule set is the following:

$$f(x, a) \in (34.4, 43.2) \Rightarrow f(x, d) = 1, (x_1)$$

$$f(x, a) \in (122.4, 156) \wedge f(x, b) = 0 \Rightarrow f(x, d) = 1, (x_7)$$

$$f(x, b) = 2 \Rightarrow f(x, d) = 2, (x_1, x_4, x_{10})$$

$$f(x, a) \in (51.6, 64.8) \Rightarrow f(x, d) = 1 \wedge f(x, d) = 2, (x_2, x_{11})$$

$$f(x, b) = 1 \Rightarrow f(x, d) = 1 \wedge f(x, d) = 2, (x_3, x_4, x_5, x_9, x_{12})$$

VI. CONCLUSIONS

When extending the rough set theory to the use of similarity relations or tolerance relations which are not necessarily symmetric and transitive, the concept of definability does not correctly capture the presence or absence of ambiguity. Therefore, we proposed the new definitions of lower and upper approximations based on the binary relation and its inverse relation. These new definitions properly characterize the set of positive objects and the set of ambiguous objects when the binary relation is not necessarily symmetric or transitive.

In order to induce rule set from a decision table involving qualitative and quantitative attributes, the nonsimilarity matrix and the nonsimilarity function are presented based on the concept of the binary relations defined from some similarity measures.

In comparison with classical approaches, the extended rough set approach to rule induction has some main advantages. First, the method of construction of a similarity measure overcomes the strict conjunctive operator traditionally used in rough set theory based on indiscernibility relation. Secondly, the decision rules induced from similarity relation are more robust than those induced from indiscernibility relation since the former are less sensitive to small differences of attribute values. Lastly, the method does not require any discretization of the domains of quantitative attributes.

ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China (No. 70601013).

REFERENCES


*Liping An*, born in 1971, received the BS degree in management science and engineering from Hebei University of Technology, Tianjin, China, in 1993, the MS degree in management science and engineering from Hebei University of Technology, Tianjin, China, in 1999, and the PhD degree in management science and engineering from Tianjin University, Tianjin, China, in 2002. Currently, he is an associate professor in the Department of Management Science and Engineering at Nankai University. He is also the vice director of the department. His research interests include business intelligence, rough set, and information system development.

*Lingyun Tong*, born in 1971, received the BS degree in management science and engineering from Hebei University of Technology, Tianjin, China, in 1993, the MS degree in management science and engineering from Hebei University of Technology, Tianjin, China, in 1996, and the PhD degree in control theory and control engineering from Nankai University, Tianjin, China, in 2005. Currently, she is an associate professor in the Department of Information System and Information Management at Hebei University of Technology. Her research interests include artificial intelligence, rough set, and supply chain management.