Classification System Based on Non-Revision Reasoning and Rough Set Theory

Zhang Song, Deng Ansheng, and Qu Yanpeng

***Abstract*—Rough set theory is an effective mathematical tool to process inaccurate, inconsistent and incomplete information. The primary goal of rough set theory has been outlined as a classificatory analysis of data: given a data table, rough set algorithms induce a set of relevant concepts such as rules providing a classification of data. However, these concepts may contradict with some priori knowledge in expert system which causes problem in reasoning. This paper proposes a classification system based on rough set theory and non-revision reasoning which tolerates the inconsistency between the generated concepts and priori knowledge. This approach integrates knowledge from multi-sources without data normalizing, which improves the efficiency and the rationality of the classification result. Moreover, integration of knowledge instead of data also preserves the information privacy and reduces the cost on transfer.**

***Index Terms*—Classification, decision rule, non-revision, rough set.**

I. Introduction

The primary goal of rough set theory has been outlined as a classificatory analysis of data: given a data table, rough set algorithms induce a set of relevant concepts providing a classification of data [1]. It is formulated in terms of indiscernibility relations and derived notions of a reduct, a decision rule and a decision algorithm used to classify. Indiscernibility relations are induced from data table. It generates equivalence class of the objects and allows for information reduction by selecting sets of attributes of the objects preserving classification. The minimal sets of these attributes are called reduct. Given a reduct, decision rules can be induced from it abstracted from attribute-value descriptors provided by the table [1]. Conventional classification system based on rough set theory uses these decision rules to decide which class a new object belongs to.

However, in practical application, priori knowledge from the expert system is necessary for classification. Rules generated by rough set method may contradict with these knowledge. In classical logic, if the premise is inconsistent, all the derivations from it are meaningless. To solve the problem, this paper proposed a non-revision reasoning [2] in classification system drawing the conclusions without any derivable reverse. The reasoning generates a unique, consistent and closed extension of the knowledge base constituted by facts of new object, priori knowledge from expert system and decision rules generated from multi-sources by rough set method. Compared with conventional classification system, it improves the rationality of the classification result for the application of priori knowledge and multi-sources decision rules. Data from different sources needn't to be normalized. Moreover, integration of decision rules instead of data also preserves the information privacy and reduces the cost on transfer.

The rest of the paper is organized as follows. In Section II, we briefly introduce the rough set theory. In Section III, we study the non-revision reasoning. In Section IV, we construct the classification system based on non-revision reasoning and rough set theory. In Section V, we give some algorithms for the system. Then we introduce a few related works in Section VI and summarize in the last section.

Manuscript received June 20, 2015; revised January 14, 2016. This work was supported by the national Natural Science Foundation of China (61272171; 61371090) and the Fundamental Research Funds for the Central Universities (NO. 3132014094) and the China Postdoctoral Science Foundation (NO. 2013M541213, 2015T80239) and the Fundamental Research Funds for the Central Universities of China (NO. 3132015044).

The authors are with Dalian Maritime University Dalian 116026, PRC (e-mail: zhangsongtt@163.com, ashdeng@dlmu.edu.cn, yanpengqu@dlmu.edu.cn).

II. Rough Set Theory

TABLE I: Decision System of Cars

Car	Comfort	Price	Color	Sales
x_1	high	high	red	normal
x_2	low	low	blue	normal
x_3	normal	low	red	low
x_4	high	normal	blue	high
x_5	high	high	red	low

Classical rough set theory is formulated in terms of indiscernibility relations and derived notions of a reduct, a decision rule and a decision algorithm used to classify. In this section, we introduce rough set theory in this order. First, we give the formal definition of decision system. Second, introduce the basic assumption of rough set theory called indiscernibility relations. Then give the definition of reduct allowing for information reduction. At last, decision rule of classification is discussed.

A. Decision System

Definition 1. Decision system is a triple $A_d = \langle U, A, d \rangle$, where U is a non-empty finite set of objects called the universe of A_d and A is a non-empty finite set of attributes and d is a distinguished attribute called decision. Any attribute a is a map $a: U \rightarrow V_a$. The set V_a is called the value set of a. The same as any other attribute, d is a map $d: U \to V_d$. And $\mathrm{A} = \langle U, A \rangle$ is called information system.

A decision system can be expressed as an attribute-value data table. For example, Table I is a decision system recording the information of secondhand cars. The cars are objects. Properties like comfort, price and color are attributes. Sales volume is decision. As is known, there are some dependency relations between the properties of the car and the sales volume. The goal of classification is to discover these relations and use them to classify new object.

B. Indiscernibility Relation

Definition 2. For an object $x \in U$, the information about x borne by the information system A with respect to a set $B \subseteq A$ of attributes is defined as the B-information set

$$Inf_B(x) = \{(a, a(x)) : a \in B\}.$$

Definition 3. The relation IND_B of the B-indiscernibility is defined as follows

$$(x, y) \in IND_B \Leftrightarrow Inf_B(x) = Inf_B(y).$$

Equivalence classes $[x]_B$ of the relation IND_B represent elementary portions of knowledge represented by the subsystem $A_B = \langle U, B \rangle$, for any $B \subseteq A$.

Definition 4. Assuming that $X \subseteq U$, for any attribute set B, define the B-lower and B-upper approximation of X, denoted $\underline{B}X$ and $\overline{B}X$ as follows

$$\underline{B}X = \{x \in U : [x]_B \subseteq X\}$$

and

$$\overline{B}X = \{x : [x]_B \cap X \neq \varnothing\}.$$

The set

$$BN_B(X) = \overline{B}X - \underline{B}X$$

is called the B-boundary region of X.

The decision attribute d induces the partition of the universe U into equivalence classes of the d-indiscernibility relation IND_d. Without the loss of generality, assuming that $V_d = \{1, 2, \cdots, k(d)\}$, the partition $\{X_1, X_2, \cdots, X_{k(d)}\}$ of U is obtained into decision classes.

Definition 5. Assuming that A_d is a decision system and $B \subseteq A$ is a set of attributes,

$$POS_B(d) = \{x \in U : \exists i \in \{1, 2 \cdots, k(d)\}.[x]_B \subseteq X_i\}.$$

The set $POS_B(d)$ is called the B-positive region of d.

As the above example,

$$POS_{\{\text{comfort}\}}(\text{sales}) = \{x_2, x_3\},$$

$$POS_{\{\text{price}\}}(\text{sales}) = \{x_4\},$$

$$POS_{\{\text{comfort,price}\}}(\text{sales}) = \{x_2, x_3, x_4\}.$$

C. Reduct

Definition 6. Assuming that A_d is a decision system and the set of attribute $B \subseteq A$ is called a relative reduct of A_d if

1) $POS_B(d) = POS_A(d)$.
2) B is a minimal set of attributes with respect to the property 1.

It has been proved in [3] that finding a minimal relative reduct is an NP-hard problem. However, many approaches optimize the algorithm. According to the algorithm in [4], we can get the reduct of Table I as $\{\text{comfort}, \text{price}\}$. It means that the color of cars has no influence on the sales volume. The concrete process of computing reduct is omitted in this paper because it makes no difference to non-revision reasoning. For more details, see [4]-[7].

D. Decision Rules

Definition 7. Assume that $A_d = \langle U, A, d \rangle$ is a decision system. Logic L_{A_d} is defined as: $F \in L_{A_d}$ if and only if

1) F is an elementary formula where F is of the form (a, v) where $a \in B$ for some $B \subseteq A$.
2) F is a propositional connection of elementary formulas by $\vee, \wedge, \neg$.

The semantics of L_{A_d} is defined as

1) $[(a, v)] = \{x \in U : a(x) = v\}$.
2) $[(a, v) \vee (b, w)] = [(a, v)] \cup [(b, w)]$.
3) $[(a, v) \wedge (b, w)] = [(a, v)] \cap [(b, w)]$.
4) $\neg[(a, v)] = U - [(a, v)]$.

A decision rule for A is an expression of the form $\alpha \Rightarrow (d, v)$, where α is a formula of L_{A_d} employing only elementary sub-formula of the form (a, v) with $a \in B$ for some $B \subseteq A$ and the meaning $[\alpha] \neq \varnothing$. Formula α and (d, v) are referred to as the predecessor and the successor of the decision rule $\alpha \Rightarrow (d, v)$. $\alpha \Rightarrow (d, v)$ is true in A_d if and only if $[\alpha] \subseteq [(d, v)]$.

Decision rules can be generated from the reduction of the decision system which has the same classification as it. The reduct of Table I is $\{\text{comfort}, \text{price}\}$ from which decision rules are generated as

$$(\text{comfort}, \text{low}) \wedge (\text{price}, \text{low}) \Rightarrow (\text{sales}, \text{normal}),$$

$$(\text{comfort}, \text{normal}) \wedge (\text{price}, \text{low}) \Rightarrow (\text{sales}, \text{low}),$$

$$(\text{comfort}, \text{high}) \wedge (\text{price}, \text{normal}) \Rightarrow (\text{sales}, \text{high}).$$

The accuracy of a rule is the quotient $\frac{\|[\alpha] \cap [(d, v)]\|}{\|[\alpha]\|}$ while the coverage of the rule is defined as the quotient $\frac{\|[\alpha] \cap [(d, v)]\|}{\|[(d, v)]\|}$. We can find that if $[\alpha] \subseteq \underline{B}[(d, v)]$, the accuracy of $\alpha \Rightarrow (d, v)$ is 1 while the coverage is on $(0,1)$.

If $[\alpha] \subseteq BN_B([(d,v)])$, the accuracy of $\alpha \Rightarrow (d,v)$ is on $(0,1)$ while the coverage is on $(0,1]$. Some approaches like [8]-[11] qualify the preference of decision rules by giving a quality function mapping accuracy and coverage to a numeric degree.

Decision rules with accuracy less than 1 can be generated from Table I as

$$(\text{comfort}, \text{high}) \wedge (\text{price}, \text{high}) \Rightarrow (\text{sales}, \text{normal}),$$

$$(\text{comfort}, \text{high}) \wedge (\text{price}, \text{high}) \Rightarrow (\text{sales}, \text{low}).$$

Decision rules can be used to classify new objects. For example, if x_6 is a car to be classified of which comfort is high and price is normal, it can be donated as a formula $(\text{comfort}, \text{high}) \wedge (\text{price}, \text{normal})$ of x_6. By matching it with the predecessors of the decision rules, $(\text{sales}, \text{high})$ can be derived as the result of the classification showing that the sales volume of x_6 is high.

However, in practical application, priori knowledge from expert system is necessary. These knowledge may contradict with the decision rules. In next section, we propose a non-revision reasoning tolerating the inconsistency in knowledge base. Facts of new objects to be classified, priori knowledge from expert system and decision rules generated by decision system constitute the knowledge base of it. By reasoning, classification result is obtained in the extension set of the knowledge base.

III. Non-Revision Reasoning

Non-revision reasoning proposed in [2] is an effective approach to deal with the inconsistency in knowledge-based system. This paper applies the work of [12] proposed in predicate logic. In non-revision reasoning, inaccurate and incomplete knowledge is called "belief". When the belief base is inconsistent, non-revision reasoning method draws the conclusions without any derivable reverse. It generates a unique, consistent and closed extension set of the belief base instead of revising it. In this section, a series of definitions are given for the extension set and some properties of it are raised.

A. Extension of Belief Base

As is assumed in [12], belief is a clause in predicate logic. Belief base is a clause set expressed as Γ. The reasoning follows resolution principle R [13]. All the derivations of the resolution and the belief base constitute the set $Con(\Gamma)$.

Definition 8. Belief base is a clause set in predicate logic.

Definition 9. Assuming that Γ is clause set, $Con(\Gamma)$ is defined as follow: clause $C \in Con(\Gamma)$ if and only if there exists a finite sequence of clauses $C_1, C_2, \cdots, C_n$, where

1) For any clause C_i in the sequence, either $C_i \in \Gamma$ or there are clauses C_s, C_t precede C_i in the sequence, such that $R(C_s, C_t) = C_i$.
2) $C = C_n$.

For example, assume $\Gamma = \{C, D, E\}$, where

$$C = P(x) \vee Q(x),$$

$$D = \neg P(a),$$

$$E = \neg Q(a).$$

Then there are three new beliefs can be derived from Γ:

$$F = R(C, D) = Q(a),$$

$$G = R(C, E) = P(a),$$

$$H = R(D, G) = R(E, F) = \square.$$

According to the definition,

$$Con(\Gamma) = \{C, D, E, F, G, H\}.$$

$Con(\Gamma)$ is an extension of the belief base closed on R. However, $Con(\Gamma)$ may be inconsistent because empty clause can be derived from it. Some definitions are given to deal with the inconsistency.

Definition 10. Justification j is a double $\langle Cl(j), Sup(j) \rangle$, where $Cl(j)$ is a clause in $Con(\Gamma)$ and $Sup(j)$ is a clause set containing clauses in $Con(\Gamma)$ used to derived $Cl(j)$. Call j a justification of $Cl(j)$. If C is a clause in belief base, then there must be a justification $\langle C, \varnothing \rangle$ of C.

If a belief is derived from different sources, the system produces different justifications to it. As the above example,

$$R(D, G) = \square \text{ and } R(E, F) = \square.$$

$\square$ can be derived from D, G and E, F. So the system produces two justifications to $\square$.

$$\langle \square, \{D, G\} \rangle \text{ and } \langle \square, \{E, F\} \rangle.$$

Beliefs in belief base are initialized before reasoning, so each of them has a justification with empty support set. In this example, $C \in \Gamma$. So C has a justification

$$\langle P(x) \vee Q(x), \varnothing \rangle.$$

Definition 11. Assuming that C is a clause in $Con(\Gamma)$ and j is a justification, $Pre(C)$ and $Qua(j)$ are numerical preferences on $[0,1]$. And define $Pre(C)$ as

$$Pre(C) = \max\{Qua(j_1), Qua(j_2), \cdots, Qua(j_n)\},$$

where $j_1, j_2, \cdots, j_n$ are all the justifications of C. And define $Qua(j)$ as

$$Qua(j) = \min\{Pre(C) : C \in Sup(j)\}.$$

As the above example, $Con(\Gamma) = \{C, D, E, F, G, H\}$, the Qua values of the initial justifications are given as

$$j_C = \langle P(x) \vee Q(x), \varnothing \rangle \ \ 0.7,$$

$$j_D = \langle \neg P(a), \varnothing \rangle \ \ 0.5,$$

$$j_E = \langle \neg Q(a), \varnothing \rangle \ \ 0.4,$$

After the reasoning, the preference are modified as

$$C: \ \max\{Qua(j_C)\} = 0.7,$$

$$D: \ \max\{Qua(j_D)\} = 0.5,$$

$$E: \ \max\{Qua(j_E)\} = 0.4,$$

$$j_F = \langle Q(a), \{C, D\} \rangle: \ \min\{Pre(C), Pre(D)\} = 0.5,$$

$$F: \ \max\{Qua(j_F)\} = 0.5,$$

$$j_G = \langle P(a), \{C, E\} \rangle: \ \min\{Pre(C), Pre(E)\} = 0.4,$$

$$G: \ \max\{Qua(j_G)\} = 0.4,$$

$$j_{H1} = \langle \square, \{D, G\} \rangle: \ \min\{Pre(D), Pre(G)\} = 0.4,$$

$$j_{H2} = \langle \square, \{E, F\} \rangle: \ \min\{Pre(E), Pre(F)\} = 0.4,$$

$$H: \ \max\{Qua(j_{H1}), Qua(j_{H2})\} = 0.4.$$

Definition 12. Assuming that Γ is a clause set, $Min(\Gamma)$ is defined as follow: clause $C \in Min(\Gamma)$ if and only if $C \in \Gamma$ and for any clause $D \in \Gamma$, $Pre(C) \le Pre(D)$.

For example, if $\Gamma = \{C, D\}$, where

$$C = P(x): \ 0.4,$$

$$D = \neg P(a): \ 0.7.$$

Then $Min(\Gamma) = \{C\}$.

And if $\Gamma = \{C, D\}$, where

$$C = P(x): \ 0.4,$$

$$D = \neg P(a): \ 0.4.$$

Then $Min(\Gamma) = \{C, D\}$.

Definition 13. Assuming that Γ is a clause set, $Del(\Gamma)$ is a set defined as follows: clause $C \in Del(\Gamma)$ if and only if either $C \in Con(\Gamma)$ and $C = \square$, or there exists a clause D such that $D \in Del(\Gamma)$ and $C \in Min(Sup(j))$ where j is a justification of D.

Such as $Con(\Gamma) = \{C, D, E, F, G, H\}$ in above example,

$$H = \square$$

$$j_{H1} = \langle \square, \{D, G\} \rangle \ \ 0.4,$$

$$j_{H2} = \langle \square, \{E, F\} \rangle \ \ 0.4,$$

$$\mathrm{Min}(Sup(j_{H1})) = \mathrm{Min}(\{D, G\}) = \{G\},$$

$$\mathrm{Min}(Sup(j_{H2})) = \mathrm{Min}(\{E, F\}) = \{E\},$$

$$j_G = \langle P(a), \{C, E\} \rangle,$$

$$\mathrm{Min}(Sup(j_G)) = \mathrm{Min}(\{C, E\}) = \{E\},$$

$$j_E = \langle \neg Q(a), \varnothing \rangle,$$

$$\mathrm{Min}(Sup(j_E)) = \varnothing.$$

According to the definition, $Del(\Gamma) = \{H, G, E\}$.

Definition 14. Assuming that Γ is a clause set, define $R(\Gamma)$ as

$$R(\Gamma) = Con(\Gamma) - Del(\Gamma).$$

B. Properties of Extension

$R(\Gamma)$ is the extension set of belief base. It has been proved in [12] that $R(\Gamma)$ is unique, closed, consistent, cumulative and stable. The epistemic process [14] of $R(\Gamma)$ is convergent. Assuming that Γ is a belief base, cumulativity and stability can be expressed as follows.

Cumulativity: $C \in R(\Gamma) \Rightarrow R(\Gamma) = R(\Gamma \cup C)$.

Stability: $C \in R(\Gamma) \Rightarrow C \in R(\Gamma \cup C)$.

IV. Classification System

The term "classification" concerns any context in which some decision is taken or a forecast is made on the basis of currently available knowledge or information [1]. This paper proposed a classification system based on non-revision reasoning and rough set theory. Facts of new objects to be classified, priori knowledge and decision rules generated by rough set method constitute the belief base of non-revision reasoning. It generates a unique, consistent and closed extension set of belief base. Beliefs in the extension set represented a final or intermediate result of the classification. The first step to construct the classification system is to translate L_{A_d} logic to predicate logic. Second, we need to provide quality function of decision rules to denote the preference of beliefs in non-revision reasoning. Then formulize facts of new object to be classified into single literal clause and add them to belief base. After reasoning, derivations with single literal in extension set are selected as the classification result. If decision rules are generated from multi-sources, the reliability of the sources should be considered in the quality function and rules from each source are added into belief base after translation.

A. Translation

According to the definition in Section II, decision rule generated by rough set method is an expression of the form $\alpha \Rightarrow (d, v)$ where α is represented in logic L_{A_d}, while belief in non-revision reasoning is represented as clause in predicate logic. The translation method is firstly to assume predicates with the same semantics as L_{A_d} and replace formulas of it. Then transform these rules into clauses by equivalence. This can be formalized as follows.

For a decision system $A_d = \langle U, A, d \rangle$, decision rule $\text{Rule} = (\alpha \Rightarrow (d, v))$ generated from A_d where α is a formula of L_{A_d} can be expressed without the loss of generality as

$$\alpha = (a_1, v_1) \wedge \cdots \wedge (a_k, v_k)$$

where $a_1, \cdots, a_k \in A$ and $v_i \in V_{a_i}, (0 < i \le k)$.

The translation is a process from Rule to clause $Cl(\text{Rule})$ by

1) Replacing (a_i, v_i) by predicate $r_{(a_i,v_i)}(x)$ where x is a variable on U. $r_{(a_i,v_i)}(x)$ is true if and only if $a_i(x) = v_i$.
2) Replacing (d, v) by predicate $p_v(x)$ where x is a variable on U. $p_v(x)$ is true if and only if $d(x) = v$.
3) Replacing $\Rightarrow$ by $\rightarrow$.
4) Transforming the rule into clause by equivalence.

As the example in Section II, the decision rules generated from Table I can be translated as

$$\text{Rule}_1 = ((\text{comfort}, \text{low}) \wedge (\text{price}, \text{low}) \Rightarrow (\text{sales}, \text{normal})),$$

$$\text{Rule}_2 = ((\text{comfort}, \text{normal}) \wedge (\text{price}, \text{low}) \Rightarrow (\text{sales}, \text{low})),$$

$$\text{Rule}_3 = ((\text{comfort}, \text{high}) \wedge (\text{price}, \text{normal}) \Rightarrow (\text{sales}, \text{high})),$$

$$\text{Rule}_4 = ((\text{comfort}, \text{high}) \wedge (\text{price}, \text{high}) \Rightarrow (\text{sales}, \text{normal})),$$

$$\text{Rule}_5 = ((\text{comfort}, \text{high}) \wedge (\text{price}, \text{high}) \Rightarrow (\text{sales}, \text{low})).$$

$$Cl(\text{Rule}_1) = \neg r_{(\text{comfort,low})}(x) \vee \neg r_{(\text{price,low})}(x) \vee p_{\text{normal}}(x),$$

$$Cl(\text{Rule}_2) = \neg r_{(\text{comfort,normal})}(x) \vee \neg r_{(\text{price,low})}(x) \vee p_{\text{low}}(x),$$

$$Cl(\text{Rule}_3) = \neg r_{(\text{comfort,high})}(x) \vee \neg r_{(\text{price,normal})}(x) \vee p_{\text{high}}(x),$$

$$Cl(\text{Rule}_4) = \neg r_{(\text{comfort,high})}(x) \vee \neg r_{(\text{price,high})}(x) \vee p_{\text{normal}}(x),$$

$$Cl(\text{Rule}_5) = \neg r_{(\text{comfort,high})}(x) \vee \neg r_{(\text{price,high})}(x) \vee p_{\text{low}}(x).$$

B. Quality Function

In non-revision reasoning, the quality of a justification with empty support set should be initialized before reasoning. To achieve this, we give the quality function which maps the coverage and accuracy of the decision rule to a numeric degree.

Definition 15. Assume that rule is a decision rule generated from decision system $A_d = \langle U, A, d \rangle$. Quality function f is a function from $[0,1] \times [0,1]$ to $[0,1]$.

Let $Qua(\langle Cl(Rule), \varnothing \rangle) = f(\varepsilon, \eta)$, where

$$\varepsilon = \text{coverage}_{A_d}(\text{Rule}),$$

$$\eta = \text{accuracy}_{A_d}(\text{Rule}).$$

Some quality functions have been defined in [8]-[11]. We apply the function $f(\varepsilon, \eta) = \omega \cdot \varepsilon + (1 - \omega) \cdot \eta$ where $\omega = 0.5$ in this paper. The quality of justifications generated from Table I is

$$Qua(\langle Cl(\text{Rule}_1), \varnothing \rangle) = 0.75,$$

$$Qua(\langle Cl(\text{Rule}_2), \varnothing \rangle) = 0.75,$$

$$Qua(\langle Cl(\text{Rule}_3), \varnothing \rangle) = 1,$$

$$Qua(\langle Cl(\text{Rule}_4), \varnothing \rangle) = 0.5,$$

$$Qua(\langle Cl(\text{Rule}_5), \varnothing \rangle) = 0.5.$$

C. Classification

Definition 16. Assuming that clause $C \in \Gamma$, $|\text{literal}(C)|$ is the number of literals in C. Define $\text{Fact}(\Gamma)$ as follow: clause $C \in \text{Fact}(\Gamma)$ if and only if $|\text{literal}(C)| = 1$. Define $\text{Rule}(\Gamma)$ as follow: clause $C \in \text{Rule}(\Gamma)$ if and only if $|\text{literal}(C)| > 1$.

Definition 17. Assuming that Γ is a clause set, $C \in \text{Max}(\Gamma)$ is defined as follow: clause $C \in \text{Max}(\Gamma)$ if and only if $C \in \Gamma$ and for any clause $D \in \Gamma$, $Pre(C) > Pre(D)$.

Before classification, facts of new object to be classified are added into the belief base. It is represented by the clause with single literal recording the information of attribute-value of the new object. For example, if x_6 is a car to be classified of which comfort is high and price is normal, it can be denoted as a set of clauses

$$\Gamma_{\text{Fact}} = \{ r_{(\text{comfort,high})}(x_6), r_{(\text{price,normal})}(x_6) \}.$$

Assuming that the clause set translated from decision rules is Γ_{Rule} and priori knowledge set from expert system is Γ_{Expert}, let belief base $\Gamma = \Gamma_{\text{Fact}} \cup \Gamma_{\text{Rule}} \cup \Gamma_{\text{Expert}}$. After reasoning, $R(\Gamma)$ is derived. Clauses in

$$\text{Max}(\text{Fact}(R(\Gamma) - \Gamma))$$

represent the final results of classification and clauses in

$$\text{Rule}(R(\Gamma) - \Gamma)$$

represent intermediate results.

For example of Table I,

$$\Gamma_{\text{Rule}} = \{ Cl(\text{Rule}_1), Cl(\text{Rule}_2), Cl(\text{Rule}_3), Cl(\text{Rule}_4), Cl(\text{Rule}_5) \}.$$

Assume that $\Gamma_{\text{Expert}} = \{\neg r_{(\text{wear,high})}(x) \vee \neg p_{\text{high}}(x)\}$ where predicate $r_{(\text{wear,high})}(x)$ means that the degree of wear of a car is high.

Assume that

$$\Gamma_{\text{Fact}} = \{C, D, E, F\}$$

where

$$C = r_{(\text{comfort,high})}(x_6),$$

$$D = r_{(\text{price,normal})}(x_6),$$

$$E = r_{(\text{color,red})}(x_6),$$

$$F = r_{(\text{wear,high})}(x_6).$$

It means that the new object to be classified is car x_6 with high comfort, normal price, red color and high wear. Let $\Gamma = \Gamma_{\text{Fact}} \cup \Gamma_{\text{Rule}} \cup \Gamma_{\text{Expert}}$ and the quality of justifications is

$$Qua(\langle Cl(\text{Rule}_1), \varnothing \rangle) = 0.75,$$

$$Qua(\langle Cl(\text{Rule}_2), \varnothing \rangle) = 0.75,$$

$$Qua(\langle Cl(\text{Rule}_3), \varnothing \rangle) = 1,$$

$$Qua(\langle Cl(\text{Rule}_4), \varnothing \rangle) = 0.5,$$

$$Qua(\langle Cl(\text{Rule}_5), \varnothing \rangle) = 0.5.$$

$$Qua(\langle \neg r_{(\text{wear,normal})}(x) \vee \neg p_{\text{high}}(x), \varnothing \rangle) = 0.8,$$

$$Qua(\langle C, \varnothing \rangle) = 1,$$

$$Qua(\langle D, \varnothing \rangle) = 0.75,$$

$$Qua(\langle E, \varnothing \rangle) = 0.75,$$

$$Qua(\langle F, \varnothing \rangle) = 1.$$

After reasoning,

$$\text{Fact}(R(\Gamma) - \Gamma) = \{\neg p_{\text{high}}(x_6)\},$$

$$Max(\text{Fact}(R(\Gamma) - \Gamma)) = \{\neg p_{\text{high}}(x_6)\}.$$

So the final result of classification is $\{\neg p_{\text{high}}(x_6)\}$ which means the sales volume of car x_6 is not high.

D. Multi-sources

If the decision rules are generated from multi-sources, some factors of the sources like data size, achieving approach and so on make a difference on reliability of the rules. It should be considered in the quality function.

Assume that $\{A_{d1}, \cdots, A_{dn}\}$ are a set of decision systems and $\Gamma_{\text{Rule}i}$ is a clause set generated from A_{di} where $0 < i \le n$. $\text{reliability}(A_{di})$ is a number on $[0,1]$ borne the reliability of A_{di}. The multi-sources quality function f_{Multi} is a function from $[0,1] \times [0,1] \times [0,1]$ to $[0,1]$.

For $\text{Rule} \in \Gamma_{\text{Rule}i}$, let

$$Qua(\langle Cl(\text{Rule}), \varnothing \rangle) = f_{\text{Multi}}(\varepsilon, \eta, \mu),$$

where

$$\varepsilon = \text{coverage}_{A_{di}}(\text{Rule}),$$

$$\eta = \text{accuracy}_{A_{di}}(\text{Rule}),$$

$$\mu = \text{reliability}(A_{di}).$$

Let $\Gamma = \Gamma_{Fact} \cup \Gamma_{Expert} \cup \Gamma_{Rulei} \cup \cdots \cup \Gamma_{Rulen}$.

After reasoning, the final result of the classification for multi-sources is in $Max(Fact(R(\Gamma) - \Gamma))$.

V. Complexity

In this section, we give an algorithm of the classification system. In algorithm 1, function *Reduct* computes the reduct of the decision system. Its complexity has been discussed in [15], [16]. Function *Generate* generates the decision rules to the set *DecisionRules*. Function *Translate* translates the decision rules to clause set Γ_{Rule}. Function *Quality* computes the *Qua* of the justification of clauses in Γ_{Rule} by quality function f. They all run in polynomial time by the size of the reduct. Function $R(\Gamma)$ derives the extension of Γ which runs in time $O(kn)$, where n is the size of $Con(\Gamma)$ and k ranges $[1, n]$ according to the good or bad situation. As for $Con(\Gamma)$, there is no efficient algorithm running in polynomial time because of the complexity of resolution reasoning. However, in [12], a threshold λ is used to reduce the complexity.

Assuming λ is a real number on $[0,1]$, we define Γ^{λ} as follow: belief $C \in \Gamma^{\lambda}$ if and only if $C \in \Gamma$ and $Qua(\langle C, \varnothing \rangle) \ge \lambda$.

When there is not enough time to compute the whole $Con(\Gamma)$, $Con(\Gamma^{\lambda})$ can be a second choice. Each belief in $Con(\Gamma^{\lambda})$ has a degree of preference no less than λ which is more preferred relatively. The value of λ can be adjusted according to the requirement of time or preference to adapt to different application environment.

Algorithm 1. **Classification system**

Input decision system A_d

Input clause set Γ_{Fact} and *Qua* of justifications

Input priori knowledge set Γ_{Expert} and *Qua* of justifications

$\text{Reduct}(A_d)$

$\text{DecisonRules} = \text{Generate}(A_d)$

$\Gamma_{\text{Rule}} = \text{Translate}(\text{DecisionRules})$

$\text{Quality}(\Gamma_{\text{Rule}})$

$\Gamma = \Gamma_{\text{Fact}} \cup \Gamma_{\text{Rule}} \cup \Gamma_{\text{Expert}}$

$R(\Gamma)$

Output $\text{Max}(\text{Fact}(R(\Gamma) - \Gamma))$

For multi-sources classification system, clause sets $\Gamma_{\text{Rule}i}, \cdots, \Gamma_{\text{Rule}n}$ are generated by different agents. Function $\text{Quality}_{\text{Multi}}(\Gamma_{\text{Rule}i})$ computes Qua of justification of clauses in $\Gamma_{\text{Rule}i}$ by multi-sources quality function f_{Multi}. After reasoning, the final result of the classification for multi-sources is in $\text{Max}(\text{Fact}(R(\Gamma) - \Gamma))$.

Algorithm 2. **Multi-sources classification system**

Input clause sets $\Gamma_{\text{Rule}i}, \cdots, \Gamma_{\text{Rule}n}$ generated from multi-sources

Input clause set Γ_{Fact} and Qua of justifications

Input priori knowledge set Γ_{Expert} and Qua of justifications

For $0 < i \le n$ **do**

$\text{Quality}_{\text{Multi}}(\Gamma_{\text{Rule}i})$

End for

$\Gamma = \Gamma_{\text{Fact}} \cup \Gamma_{\text{Expert}} \cup \Gamma_{\text{Rule}i} \cup \cdots \cup \Gamma_{\text{Rule}n}$

$R(\Gamma)$

Output $\text{Max}(\text{Fact}(R(\Gamma) - \Gamma))$

VI. Related Works

Some related approaches on classification such as [17], [18] usually concentrate on reduct algorithm of rough set theory, which do not take priori knowledge into consideration. In real application, priori knowledge plays an important role for covering the shortage of rough set method which is affected by noise or incomplete information. As the above example, decision rules generated from Table I reflex the relationship between sales volume and properties like comfort and price of the secondhand cars. However, the degree of wear is another critical factor on sales volume which is not considered in Table I. Priori knowledge extends the decision system for classification. It improves the rationality of the result.

Approaches based on rough set theory for multi-agents like [19], [20] mainly deal with the problem on sharing data between agents. This paper proposes an approach of sharing decision rules between agents. Decision rules are general and concise without detailed information. They needn't to be normalized, although they may be generated from different decision table with different attributes obtained from different agents. And the integration of decision rules instead of data also preserves the information privacy and reduces the transfer cost of the systems based on mass data.

VII. Conclusion

This paper proposes a classification system based on non-revision reasoning and rough set theory. It generates decision rules for classification from decision table and integrates them with priori knowledge from expert system. When inconsistency is discovered, non-revision reasoning draws the conclusions without any derivable reverse and constitutes a unique, consistent and closed extension set containing the classification result. This approach integrates knowledge from multi-sources without data normalizing, which improves the efficiency and the rationality of the system. Moreover, integration of knowledge instead of data also preserves the information privacy and reduces the cost on transfer.

Acknowledgment

Zhang Song thanks Dr. Deng Ansheng and the members of his laboratory for their invaluable cooperation in preparing this research.

References

[1] A. Cichocki and R. Unbehaven, *Studies in Fuzziness and Soft Computing*, Physical-Verlag, 1992.

[2] D. Ansheng, "An up to down reasoning for inconsistent beliefs," *Journal of XiaMen University: Natural Science*, vol. 44, pp. 222-224, 2005.

[3] S. H. Nguyen and A. Skowron, "Quantization of real value attributes-rough set and boolean reasoning approach," in *Proc. the Second Joint Annual Conference on Information Sciences*, Wrightsville Beach, North Carolina, Sept. 28-Oct. 1, 1995.

[4] J. Zhang *et al*., "A new heuristic reduct algorithm base on rough sets theory," *Advances in Web-Age Information Management*, Springer Berlin Heidelberg, 2003.

[5] H. Keyun *et al.*, "A heuristic optimal reduct algorithm," *Intelligent Data Engineering and Automated Learning*, 2000.

[6] D. M. Zhao, H. Yue, and G. Shu, "New reduct algorithm of decision table for electric power grid fault diagnosis," *Automation of Electric Power Systems*, vol. 28, no. 4, 2004.

[7] C. Velayutham and K. Thangavel, "Unsupervised quick reduct algorithm using rough set theory," *Journal of Electronic Science and Technology*, vol. 9, pp. 193-201, 2011.

[8] Y. L. Li *et al*., "An integrated method of rough set, Kano's model and AHP for rating customer requirements' final importance," *Expert Systems with Applications*, vol. 36, no. 3, pp. 7045-7053, 2009.

[9] Y.-L. Li *et al*., "A rough set approach for estimating correlation measures in quality function deployment," *Information Sciences*, vol. 189, no. 7, pp. 126-142, 2012.

[10] Y. Ming *et al*., "A rough set-based quality function deployment (QFD) approach for environmental performance evaluation: a case of offshore oil and gas operations," *Journal of Cleaner Production*, vol. 19, no. 13, pp. 1513-1526, 2011.

[11] L.-Y. Zhai, L.-P. Khoo, and Z.-W. Zhong, "A rough set enhanced fuzzy approach to quality function deployment.," *The International Journal of Advanced Manufacturing Technology*, vol. 37, no. 5-6, pp. 613-624, 2008.

[12] S. Zhang and A. S. Deng, "Non-revision reasoning without redundancy in predicate logic," *FCST*, 2015.

[13] J. A. Robinson, "A machine-oriented logic based on the resolution principle," *Journal of the ACM (JACM)*, vol. 12, no. 1, pp. 23-41, 1965.

[14] L. I. Wei, "An open logic system," *Science in China*, vol. 36, no. 3, pp. 362-375, 1993.

[15] L. Polkowski, *Rough Sets: Mathematical Foundations*, Springer Science & Business Media, 2013.

[16] B. X. Yue *et al*., "A new rough set reduct algorithm based on particle swarm optimization," *Bio-inspired Modeling of Cognitive Tasks*, Springer Berlin Heidelberg, 2007.

[17] J. G. Bazan *et al*., "Rough set algorithms in classification problem," *Rough Set Methods and Applications*, Physica-Verlag HD, 2000.

[18] L. Yee *et al*., "A rough set approach for the discovery of classification rules in interval-valued information systems," *International Journal of Approximate Reasoning*, vol. 47, no. 2, pp. 233-246, 2008.

[19] L. T. Yen and N. Cercone, *Rough Sets and Data Mining: Analysis of Imprecise Data*, Springer Science & Business Media, 2012.

[20] S. H. Nguyen, A. Skowron, and J. Stepaniuk, “Granular computing: A rough set approach,” *Computational Intelligence*, vol. 17, no. 3, pp. 514-544, 2001.

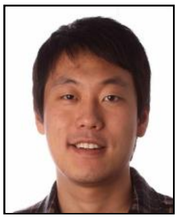
Zhang Song was born in Liaoning Province, China, in 1990. He received the B.S. degree of information science and technology from Dalian Maritime University, Liaoning, in 2013.

He is studying for the M.S. degree of information science and technology in Dalian Maritime University, Liaoning. His accepted articles are non-revision reasoning with beliefs in predicate logic (Zhangjiajie, Hunan: FSKD, 2015), non-revision reasoning without redundancy in predicate logic (Dalian, Liaoning: FCST, 2015). His research interests include automated reasoning and machine learning.

Deng Ansheng was born in June 1963, and bred in Jiaohe city, Jilin Province of the P. R. China. He received his B.S. from Beijing University in 1985, his M.S. and Ph.D. in computer science from Jilin University in 1988 and 1995, respectively.

He was with the Department of Computer Science at Northeast Normal University from June 1988 to May 2001, and the School of Information Science and Technology at Xiamen University from June 2001 to July 2005, then joining the School of Computer and Information Technology at Beijing Jiaotong University until December 2006. Currently he is a professor at the School of Information Science and Technology of Dalian Maritime University.

Dr. Deng's research is in the area of artificial intelligence and computer algorithm. His research deals mainly with logic foundations of artificial intelligence and automated reasoning in classical or non-classical logic. He has involved with 8 projects supported by the national Natural Science Foundation of China, and 2 projects supported by the National High Technique Developing Program. He has published over 60 research papers.



Qu Yanpeng is a lecturer with the Information Science and Technology College at Dalian Maritime University, China. He received a PhD degree in computational mathematics from Dalian University of Technology, China. His current research interests include rough and fuzzy set theory, pattern recognition, neural networks, classification and feature selection.