Methodology of Data Mining – Utilization of Ruduct and Indentification of Decision Rule -

Kaede Niwano[1] and Kyoichi Kijima[2]
[1]Graduate School of Decision Science and Technology, Tokyo Institute of Technology
2-12-1, Ookayama, Meguro-Ku, Tokyo, Japan
kaede@valdes.titech.ac.jp
[2]Graduate School of Decision Science and Technology, Tokyo Institute of Technology
2-12-1, Ookayama, Meguro-Ku, Tokyo, Japan
kijima@valdes.titech.ac.at

ABSTRACT

This paper proposes a new data mining methodology of reasoning for sorting new cases by using past cases or data based on rough set theory. Rough set theory is a mathematical framework to treat uncertainty. Our methodology consists of two parts. At first we focus on finding reduct for reducing large amount of data. When we try to discover a useful knowledge from a large amount of data when performing data mining, there exists trade-off between computation capacity and precision degree of the results. If we want to get more precise knowledge, we more need to cut down the data. Reduct is such a subset of the set of data attributes that provides the same quality of reasoning as the original set. Therefore by finding out an appropriate reduct, we can cut the set of attributes.

Then we introduce a new way to identify more appropriate decision by employing concept of approximation and similarity relation *S*. Generally, in reasoning using rough set theory, we have two types of rules, namely, deterministic rule and non-deterministic rule. The former determines only one decision class while the latter does not. To treat the latter cases, we propose a new approach to making more precise decisions by measuring the possibility of being in a decision class.

Keywords: Rough set theory, Knowledge, Data mining, Decision making

1. INTRODUCTION

The purpose of the present paper is to propose a new data mining methodology of reasoning for sorting new cases by using past cases or data based on rough set theory. Rough set theory is a mathematical framework to treat uncertainty.

Now a days we in information-overflowed society constantly face the need of processing large amount of information. However, we have to compromise at some point due to the bounded rationality as H. A. Simon pointed out [1]. Therefore it is important to dig out effective (important) pieces of information from ineffective ones to sort out substantial alternatives.

The importance of such data screening has been recognized most in economics. The reserch on utilization of information has been investigated as a reserch on a data mining. Though there are many methods for data mining, for instance the statics analysis and neuralnetwork, these methods have some limits. First, it is a limit of the data volume. In statics analysis, we have to do re-sampling repeatedly to gain sumpling distribution, which costs time complexity. In addition the analysis can deal with only quantitative data. In basket analysis, we have to check many combination of items. Hence, if the methods refer to NP-complete problem, we can not compute in polynominal time. Secondary, some limits arise from the method itsself. In neuoral network, the final weight is changeable depending on initial value, and the accuracy of prediction can be change every term.

In this paper, we propose a new reasoning method for treating uncertainty in data mining. We use rough set theory, which can deal with uncertainty with high precision of reasoning as well as high capability of prediction. The theory has been used in many research on data mining these days, especially in knowledge discovery, in prediction [2][3] and so many domains[5].

we will explain about rough set theory in section 2 and about decision rule in section 3. Then, a new methodology of data mining using rough set theory is introduced in section 4. In section 5 we present the algorithm of our methodology. Finally, in section 6 we describe conclusion.

2. ROUGH SET THEORY

This section introduces general idea of rough set theory (Pawlak, 1982) for preparing our framework to deal with uncertainty. Besides rough set there are several approaches like fuzzy set theory to treat

vagueness and uncertainty. It is, however, true that the rough set theory provides more objective measure of uncertainty, so that we can evaluate quality of approximation objectively by it.

2. 1. Information system and indiscernibility relations

In rough set reasoning a real world is expressed by a table which is called information system, and then some equivalence relation and set approximation are introduced.

The formulation is as follows: Information system is defined by a table consisting of

U : a finite set of objects

Q : a finite set of attributes

V : $\bigcup_{q\in Q} V_q, V_q$: the domain of attribute q

f : $U \times Q \to V, f(x,q) \in V_q$ for every $q \in Q, x \in U$

Q consists of subset C and subset D, where C denotes a set of the conditional attributes, while D represents a set of the decision attributes. D is derived from C and expresses sorting of some objects. We call such table "Decision System".

Then, for any $B \subset C$ the indiscernibility relation is a relation defined on U by

$$IND(B) = \{(x,x') \in U \times U \mid \forall b \in B, f(x,b) = f(x',b)\}.$$

Since indiscernibility relation satisfy reflexive, symmetric and transitive, it is an equivalence relation. Let us denote an equivalence class of x by $[x]_B$.

2. 2. Set Approximation

Besides the decision attribute and indiscernibility relation, from the information system we may get decision rules like "if condisional attribute = α then decision attribute = β, where α is description of x in terms of conditional attributes while β is description of x in terms of decision attributes." However, in the case that we cannot determine a unique value of decision attribute corresponding to the conditional attritute, we may not derive any crisp decision rule. To deal with such cases, we need to treat uncertainty by set approximation defined as follows:

Let X be a subset of U, then we define

B-lower approximation of X : $\underline{B}X = \{x \mid [x]_B \subset X\}$

B-upper approximation of X : $\overline{B}X = \{x \mid [x]_B \cap X \neq \phi\}$

B-boundary of X : $BN_{B(}(X) = \overline{B}X - \underline{B}X$

If $X = \{x \mid f(x,d) = "\beta"\}$ and $BN_B(X) \neq \phi$ then X is called a rough set and in that case the decision system becomes non-deterministic. The lower approximation derives deterministic and crisp rules, while the upper approximation derives non-deterministic rules.

Let us illustratet the above by taking Example 1 (See Table 1). Suppose $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ and $Q = C \cup D, C = \{q_1, q_2\}, D = \{d\}$ of $X = \{x \mid f(x,d) = "accept"\}$. Then the indiscernibility relation becomes $x_2 IND(C) x_3, x_5 IND(C) x_7$. The situation is summarized by Fig 1.

Example 1

Table 1. Decision System

	Conditional attributes		Decision
	q1	q2	d
x1	20	50	Reject
x2	100	80	Accept
x3	100	80	Accept
x4	50	60	Reject
x5	90	90	Reject
x6	100	80	Accept
x7	90	90	Accept
x8	100	80	Reject

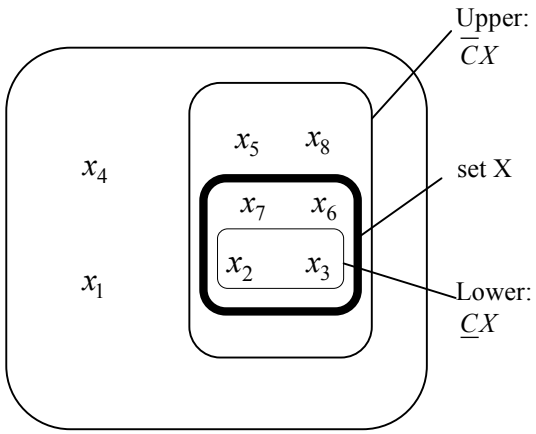


Fig 1. Set Approximations

We can have some mathematical properties for the set approximation.

$(1)\underline{B}(X) \subset X \subset \overline{B}(X)$

$(2)\underline{B}(\phi) = \overline{B}(\phi) = \phi, \underline{B}(U) = \overline{B}(U) = U$

$(3)\overline{B}(X \cup Y) = \underline{B}(X) \cup \underline{B}(Y)$

$(4)\underline{B}(X \cap Y) = \underline{B}(X) \cap \underline{B}(Y)$

$(5)X \subset Y, implies, \underline{B}(X) \subset \underline{B}(Y) and \overline{B}(X) \subset \overline{B}(Y)$

$(6)\underline{B}(X \cup Y) \supset \underline{B}(X) \cup \underline{B}(Y)$

$(7)\overline{B}(X \cap Y) \subset \overline{B}(X) \cap \overline{B}(Y)$

$(8)\underline{B}(-X) = -\overline{B}(X)$

$(9)\overline{B}(-X) = -\underline{B}(X)$

$(10)\underline{B}(\underline{B}(X)) = \overline{B}(\underline{B}(X)) = \underline{B}(X)$

$(11)\overline{B}(\overline{B}(X)) = \underline{B}(\overline{B}(X)) = \overline{B}(X)$

where $-X$ denotes $U - X$.

It is easily seen that the lower and the upper approximations of a set are, respectively, the interior and the closure of the set in terms of the topology generated by the indiscernibility relation.

2. 3. Reduct

Concept of reduct in rough set theory plays a significant and crucial role in this paper. Reduct is a subset of conditional attributes such that provides the same quality of classification as the original set of conditional attributes. Hence the ability of reduct to perform classifications is the same as that of the whole attribute set.

We define reduct in information system as follows:

R is a reduct $\Leftrightarrow R \subset C, \forall x \in U, [x]_R = [x]_c$

Then, we can introduce a decision-related reduct in such a way that it preserves the ability to identify decision. That is,

R is a decision–related reduct

$\Leftrightarrow R \subset C, \forall x \in U, [x]_R \subset [x]_D$

If decision–relative reduct R is determined, then D is automatically fixed. The fact implies that R contains sufficient infromation to identify D and we can eliminate any other information than R from the decision system for making decision on a new case.

Although reduct is very usefull concept, it is not an easy task to find the best reduct since it may be sometimes an NP-hard problem. Fortunately, however, there exist efficient heuristics based on genetic algorithms that computes sufficiently many reducts in often acceptable time, unless the number of attributes is very high. The heuristics is as follows:

Given an information system $A = (U, Q, V, f)$ with n objeccts, the discernibility matrix of the information system is a symmetric $n \times n$ matrix with entries c_{ij} defined by:

$c_{ij} = \{q \in Q \mid f(q, x_i) \neq f(q, x_j)\} for, i, j = 1, ..., n\}$

Each entry thus consists of the set of attributes upon which objects x_i and x_j differ.

A discernibility function D_A for given infromation system A is a Boolean function on m Boolean variables $q_1^*, ..., q_m^*$ (corresponding to the attributes $q_1, ..., q_m$) defined as follows.

$D_A(q_1^*, ..., q_m^*) = \wedge\{\vee c_{ij}^* \mid 1 \leq j \leq i \leq n, c_{ij} \neq \phi\}$,

where $c_{ij}^* = \{q^* \mid q \in c_{ij}\}$.

The set of all prime implicants of D_A determines the set of all reducts of A.

2. 4. Decision class

We can formalize partition of decision classes of the universe U as follows. The cardinality of the image $f(U, d) = \{k \mid f(x, d) = k, x \in X\}$ is called the rank of d and is denoted by $r(d)$. Let us assume that the set V_d of values of decision d is equal to $\{v_d^1, ..., v_d^{r(d)}\}$.

The decision d determines a partition of the universe U

$CLASS(d) = \{X_d^1, ..., X_d^{r(d)}\}$,

where $X_d^k = \{x \in X \mid d(x) = v_d^k\}$ for $1 \leq k \leq r(d)$.

Let us illustrate this using table 1 of Example1. We have image of $f(U, d) = \{accept, reject\}$,

$r(d)$=2 and

$CLASS(d) = \{X_d^1, X_d^2\}$.

The same hods for any attributes in Q.

3. DECISION RULE EXTRACTION USING ROUGH SET THEORY

We may be able to derive some decision rules from the decision system inductively using rough set theory. There are two kinds of rule. One is a “deterministic rule” derived from the lower approximation. The other is a “non-deterministic rule” obtained from the upper approximation. In Example 1, the lower and upper approximations derive the following rules, respectively.

Lower approximation : {100,80} $\Rightarrow$ accept

Upper approximation : {90,90} $\Rightarrow$ accept or reject

{100,80} $\Rightarrow$ accept or reject

In rough set theory, the description of decision attribute of a new case is derived by decision rules available so far. While the description of decision attribute of a new case is uniquely determined with a deterministic rule, it is not uniquely determined with a non-deterministic rule.

In non-deterministic case, the number of sorting examples which support each possible (decision) class is crucial. The number is called strength. If the strength of one class is greater than that of other classes one can conclude that the considered object most likely belongs

to the strongest class [5]. However, such a way of selection is too simple to describe uncertainty. We insist that we can gain more information in the decision system than that we have using another concept of approximation within rough set framework.

4. METHODOLOGY UTILIZING ROUGH SET THEORY

In this section we will propose a new approach to data mining using the rough set theory. We focus on two aspects of the theory to construct our methodology.

One of the aspects is reduct. Using it we can reduce space and time costing to treat data. The other is identification of the non-deterministic rule. We try to reason as precisely as possible by measuring the possibility of being in a decision class. We also introduce another concept of approximation to get more information that is most likely to belong to.

4. 1. Utilizing Reduct

It is important to reduce space and time costing when performing data mining. To resolve the computing resources problem we employ reduct. Since reduct is a sufficient information to make a decision rule, by using it we can save space and time by eliminating redundant attributes. We can get a reduct by performing the procedure explained in 2.3. Once we have obtained a reduct of a decision system, we only use the reduct for reasoning.

4. 2. Identification of the non-deterministic rule

We can distinguish two types of decision rule based on rough set theory, i.e., a deterministic rule and a non-deterministic rule. The deterministic rule leads to only one cleary decision class like "if a conditional attribute is α then decision attribute is β". On the other hand, the non-deterministic rule does not. It allows to more than one decision like "if a conditional attribute is α then decision attribute is β or γ ". For non-deterministic rule we generally use strength to determine which decision class a new case belongs to [5].

However, it is too simple to explain a situation. It seems inevitable way. We introduce a new approach to resolve an identification non-deterministic rule. We can devide the member of boundary of rough set into two types of member, one is positive, the other is not positive. Getting know the type, we may conclude which is more appropriate decision class by using positive or not positive.

At first, we will describe the member of positive or possible of a decision class, and then, introduce a new concept of approximation. Finally, we will get more appropriate decision class of the non-deterministic rule.

4.2.1. Positive and possible member of a decision class

When we consider a boundary on X, we introduce an original concept to set approximation [7].

Let X be a set corresponding to vague concept. Then the elements of X are not certainly in X. Elements of X can be devided into unquestionable and questionable menmers. In such a case, rough set theory can be applied to classify the elements into three categories: positive members, possible members and boundary members.

Let $\underline{X}$ and $\overline{X}$ be sets such that $\underline{X} \subset X \subset \overline{X}$. Then we call $\underline{X}$ a sets of positive members. We assume that only elements which are "similar" to a member of $\underline{X}$ can be regarded as possible members.

We call a relation S a similarity relation if it is reflexive. Let us denote

$S(x) = \{ y \mid ySx \}$,

$[x]_S = \{ y \mid xSy \}$,

$S_*(X) = \{ x \mid [x]_S \subset X \}$ and

$S^*(X) = \{ x \mid [x]_S \cap X \neq \phi \}$,

then we have

$$\overline{X} \subset \{ x \mid \exists y \in X, xSy \} = \bigcup_{y \in \underline{X}} S(y) = \{ x \mid [x]_S \cap X \neq \phi \}$$

$$\subset \{ x \mid [x]_S \cap X \neq \phi \} = S^*(X)$$

Since

$$U - \underline{X} = \overline{(U - X)} \text{ and } \underline{(U - X)} = U - \underline{X} ,$$

we also have

$$\underline{X} \supset U - \bigcup_{y \notin X} S(y) = \{ x \mid [x]_S \subset \overline{X} \}$$

$$\supset \{ x \mid [x]_S \subset X \} = S_*(X)$$

Finally, we have

$$S_*(X) \subset \underline{X} \subset X \subset \overline{X} \subset S^*(X) \qquad (1)$$

Hence we have approximations of $\underline{X}$, $S_*(X)$, and of $\overline{X}$, $S^*(X)$.

Based on the idea, we propose a new identification among elements in boundary. (1) shows that there is a difference between $\overline{X}$ and $S^*(X)$, so that we pay attention of that point. Before explaining our idea we will re-define the original notations of rough set as follows:

Lower approximation of X: $IND_*(X)$

Upper approximation of X: $IND^*(X)$

Equivalence class of x : $[x]_{IND} = \{y \mid xINDy\}$

According to considering situation, we only focus on the member of $IND^*(X)$ and since $\underline{X}$ is unquestionable members of X, we can derive

$IND^*(X) = \underline{X}$.

Hence, we have

$$\overline{X} = \{x \mid (x \in X) \cup (\exists y \in X, xSy \wedge [x]_{IND} \cap X \neq \phi)\}$$
$$\subset \{x \mid [x]_{IND} \cap X \neq \phi\} = IND^*(X)$$

and

$$\underline{X} = IND_*(X) \subset X \subset \overline{X} \subset IND^*(X).$$

This situation is shown in Fig 2. We can divide the members of $IND^*(X)$ into the members of $\overline{X}$ and not of $\overline{X}$ using above property $\overline{X} \subset IND^*(X)$. We can conclude that decision class of a x is X for all x in $IND^*(X)$, if $[x]_{IND}$ is in $\overline{X}$.

There are some types of elements in IND^* regarding relation S. We can determine a decision class of an element depends on its type.

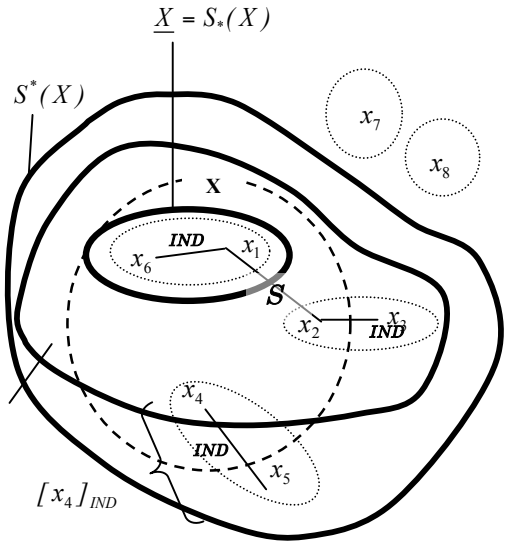
a) an element whose equivalence class is included in $\underline{X}_d^k$
b) not a) and an element with no relation S other than itself
c) not a) and b) and, an element with some relation S to the elements in same decision class
d) not a) , b) and c) and, an element with some relation S to the elements in different decision classes, the numbers of element in each decision class are different
e) not a), b), c) and d) and, an element with some relation S to the elements in different decision classes, the numbers of element in each decision class are same

We treat these types of element corresponding to above types as follows:

a) decision class is X_d^k
b) we can not determine decision class with S, re-define S and try again
c) decision class is the decision class with relation S to the element
d) decision class which have the greatest number of elements with relation S
e) we can not determine decision class with relation S, re-define S and try again

Type e) is very rare case. We had not had more precise expression of X than $\underline{X}$ and $\overline{X}$ so far. Hence re-defining S is the best way. We illustrate flow chart of identifying decision class of a new case x in Fig 3.

We may have various $\overline{X}$ depending on the definition of S. Next section (4.2.2), we discuss about a similarity relation S.



S : similarity relation
IND : indiscernibility relation

Fig 2. Image of $\underline{X}$ and $\overline{X}$

4.2.2. Similarity relation *S*

If $S(x)$ has a wide range then the member of $\overline{X}$ will increase, otherwise it will decrease. We can define a set of possible members as any set $\overline{X}$ such that $X \subset \overline{X}$, which is determined by S. Therefore, we have to consider the nature of data and define S. It is quite natural to conclude that if the description of element is similar, then these elements have the same decision class. Now we define a similarity relation S by generalizing indiscernibility relation i.e., by dropping off transitivity and symmetry from it. Consider which requirement is necessary and which is not. Reflexive is necessary since it is natural that x is a similar to x. Transitive is not necessary since it is not always true that if x is similar to y and y is similar to z then x is similar to z. Both reflexive and transitive are clear, symmetric is not. However what we shown in 4.2.1 is true even when S is not symmetric. S needs at least reflexive.

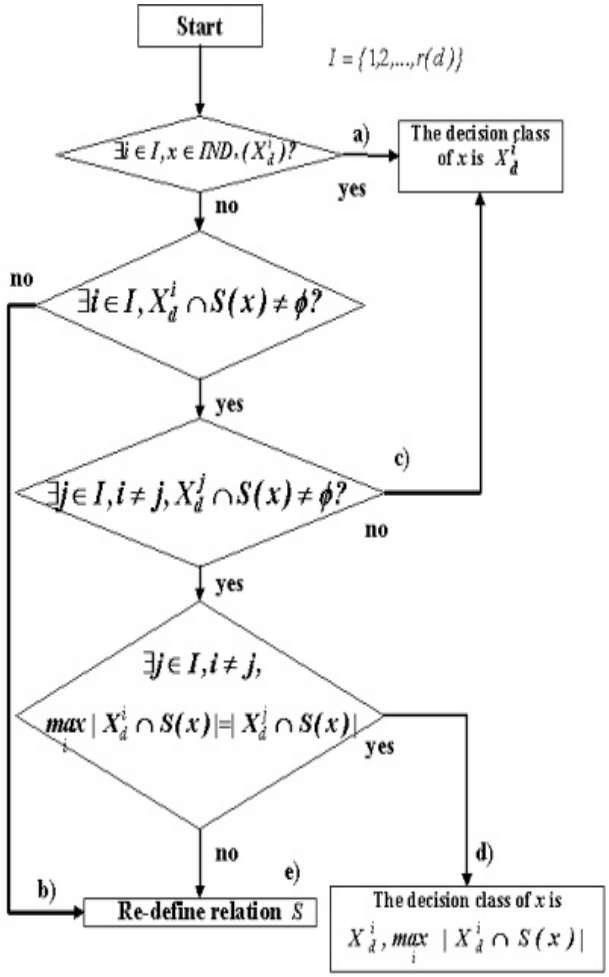


Fig 3. Flow chart of identifying decision class of x

4.2.3. Strict Similarity relation $\hat{S}$

Next we define a "strict" similarity relation $\hat{S}$ by

$\hat{S} = \{(x,x') \in U \times U \mid \forall q \in Q, d(f(x,q), f(x',q)) \leq C\}$

, where C is a constant and $d(x,x')$ shows a metric

Proposition. $\hat{S}$ is symmetric.

Proof)

We show that $(x,x') \in \hat{S} \rightarrow (x',x) \in \hat{S}$.

Suppose $(x,x') \in \hat{S}$, then

$d(f(x,q), f(x',q)) \leq C$. Since d is symmetric, we have

$d(f(x,q), f(x',q)) = d(f(x',q), f(x,q)) \leq C$

Hence, $(x',x) \in \hat{S}$. Q.E.D

It is clear when V_q is a quantitative data. We can think of an absolute value of the difference in data itself as a metric. However, in the case where V_q is a qualitative data, we cannot calculate $f(x,q)$ and $f(x',q)$. There are two ways to treat such cases in above formulation.

a) The first is to give a real number as a metric between all combination of elements in V_q.

b) The second is to give a partition of V_q

In the case a) a metric should be given as a degree of similarity. In the case b) the elements which are similar to each other should be in the same partition. If b) is adopted, $\hat{S}$ would be an equivalence relation with a broader range. Let us explain above by using Example 2. Example 2 expresses appearance of flowers, where attribute q_1 is "length of the flower" and attribute q_2 is "color of the flower". The former is quantitative while the latter is qualitative. We can calculate the metric between the elements in V_{q1} as follows:

Let $\hat{S}$ be

$\hat{S} = \{(x,x') \in U \times U \mid \forall q \in Q, d(f(x,q), f(x',q)) \leq 1\}$,

then we have

$\{x' \mid d(f(x_1,q_1), f(x',q_1) \leq 1\} = \{x_1, x_2\}$

$\{x' \mid d(f(x_2,q_1), f(x',q_1) \leq 1\} = \{x_1, x_2, x_3\}$

$\{x' \mid d(f(x_3,q_1), f(x',q_1) \leq 1\} = \{x_2, x_3, x_4\}$

$\{x' \mid d(f(x_4,q_1), f(x',q_1) \leq 1\} = \{x_3, x_4\}$

$\{x' \mid d(f(x_5,q_1), f(x',q_1) \leq 1\} = \{x_1, x_5\}$

Let us take a) for attribute q_2, and give a matrix of metric of all the combinations of V_{q2} in Table 3. Table 3 provides intuitive difference in color. Since $\hat{S}$ is a symmetric, the matrix must be a symmetric matrix. In the case of difference in color, it might be good to use RGB value depending on a considering situation.

Then we have

$\{x' \mid d(f(x_1,q_2), f(x',q_2) \leq 1\} = \{x_1, x_2, x_4\}$

$\{x' \mid d(f(x_2,q_2), f(x',q_2) \leq 1\} = \{x_1, x_2\}$

$\{x' \mid d(f(x_3,q_2), f(x',q_2) \leq 1\} = \{x_3, x_5\}$

$\{x' \mid d(f(x_4,q_2), f(x',q_2) \leq 1\} = \{x_4, x_5\}$

$\{x' \mid d(f(x_5,q_2), f(x',q_2) \leq 1\} = \{x_3, x_4, x_5\}$

Finally we have $x_1 \hat{S} x_2$ then,

$\hat{S}(x_1) = \hat{S}(x_2) = \{x_1, x_2\}$,

$\hat{S}(x_3) = \{x_3\}$, $\hat{S}(x_4) = \{x_4\}$, $\hat{S}(x_5) = \{x_5\}$,

If we take b) instead, then the partition of V_{q2} is given by

$X_{q2} = \{ X^1_{q2}, X^2_{q2} \}$,

$X^1_{q2} = \{ green, yello \}, X^2_{q2} = \{ red, blue, cyan \}$,

then

$\{ x' | d(f(x_1, q_2), f(x', q_2)) \} = \{ x_1, x_2 \}$

$\{ x' | d(f(x_3, q_2), f(x', q_2)) \} = \{ x_3, x_4, x_5 \}$.

The partition means that we treat green and yellow are the same color , and red, blue and cyan as well.

We have

$x_1 \hat{S} x_2$ and $x_3 \hat{S} x_4$ then,

$\hat{S}(x_1) = \hat{S}(x_2) = \{ x_1, x_2 \}$,

$\hat{S}(x_3) = \hat{S}(x_4) = \{ x_3, x_4 \}$,

$\hat{S}(x_5) = \{ x_5 \}$.

Example 2.

Table 2. Flower Table

	Conditional attributes	
	q_1	q_2
x_1	20.5	green
x_2	21	yellow
x_3	22.3	red
x_4	23	blue
x_5	19	cyan

Table 3. Difference in color

	green	yellow	red	blue	cyan
green	0	0.5	1.5	1	1.3
yellow	0.5	0	1.4	2	1.5
red	1.5	1.4	0	1.3	0.8
blue	1	2	1.3	0	0.8
cyan	1.3	1.5	0.8	0.8	0

5. PRESENTATION OF THE ALGORITHM

We have introduced two portions of our methodology, utilizing reduct and identification of non-deterministic rule. In this section, we present an algorithm of our methodology combining each part.

Before introducing the algorithm, we describe the repetitive processing. We must know the details of all conditional attributes in advance, when we make $\hat{S}$, otherwise we can not get reasonable result at one time. The decision class can drastically change depending on the domain of $\hat{S}$. If $\hat{S}$ has a wide domain, $\hat{S}$ could derive many relations. Unfortunately, we may not know the perfect knowledge of the conditional attributes in advance so we have to make a repetitive processing of defining relation $\hat{S}$, checking the results. Here we present the algorithm of our methodology as follows:

1) Calculate a reduct
2) Eliminate other than reduct from the whole of source data
3) Define the relation $\hat{S}$
4) Calculate the approximations of each decision class and identify decision class of all elements in upper approximation based on the flow chart in Fig. 3.
5) Check the result. If you confident of the result then finish. If not, back to 3) and repeat following procedure.

Generally, since we try to discover unknown knowledge, it seems hard to know the detailed information of attributes. It may be reasonable way to take a repetitive processing using the data.

6. CONCLUSION

In this research we applied rough set theory to data mining.

We introduced a new methodology consists of utilizing reduct and identification of the non-deterministic rule.

The former makes time- and space-complexity reduced. The data available are increasing year after year due to developing various technologies like internet. Hence we face mining problem from a large amount of data. We showed the concept of reduct is a very useful for reducing data volume.

The latter makes uncertainty reduced. Use of strength generally used in rough set theory does not take into consideration “roughness” associated with values of attributes. By introducing $\overline{X}$, similarity relation and strict similarity relation we could express rough boundaries related to decision class more flexibly.

REFERENCES

[1] H. A. Simon. (1997) “Models of Bounded Rationality”, vol.3. Empirically Grounded Economic Reason, MIT Press, 1997

[2] S. Wesley Cangchien, Tzu-Chuen Lu. (2001).”Mining association rules procedure to support on-line recommendation by customers and products fragmentation”, Expert Systems with Application, 20, 325-335.

[3] Constantin Zopounidis, Micheal Doumpos. (2002). ”Multi-criteria decision aid in financial decision making: methodologies and literature review”, of Multi-Criteria Decision Analysis, 11, 167-186.

[4] R. Slowinski, C. Zopounidis, A.I. Dimitras. (1997). “Prediction of company acquisition in Greece by means of the rough set approach”, European journal of Operational Research, 100, 1-15.

[5] Zdzislaw Pawlak. (1994). “Rough set approach to multi-attribute decision analysis”, European Journal of Operational Research, 72, 443-459.

[6] Zdzislaw Pawlak, Jerzy Grzymala-Busse, Roman Slowinski, and Wojciech Ziarko. (1995).“Rough sets”, COMMUNICATIONS OF ACM, November, Vol11, 38, 88-95.

[7] Masahiro Inuiguchi. (2001). “Generalization of Rough Sets – Rough Set based on Similarity Relation, Fuzzy Relation and Order Relation”, Japanese Journal of Fuzzy，Vol13, No.6, 562-570. (in Japanese)