Rough Sets under Non-deterministic Information

Michinori Nakata[1] and Hiroshi Sakai[2]

[1] Faculty of Management and Information Science,
Josai International University
1 Gumyo, Togane, Chiba, 283-8555, Japan
nakatam@ieee.org

[2] Department of Mathematics and Computer Aided Sciences,
Faculty of Engineering, Kyushu Institute of Technology,
Tobata, Kitakyushu, 804-8550, Japan
sakai@mns.kyutech.ac.jp

Abstract. A method of possible equivalence classes, described in [14], is extended under non-deterministic information. The method considers both indiscernibility and discernibility of non-deterministic values by using possible equivalence classes. As a result, the method gives the same results as the method of possible worlds. Furthermore, maximal possible equivalences are introduced in order to effectively calculate rough approximations. We can use the method of possible equivalence classes to obtain rough approximations between arbitrary sets of attributes containing non-deterministic values.

Keywords: Rough sets, Non-deterministic information, Imprecise values, Incomplete information, Possible equivalence classes.

1 Introduction

The framework of rough sets, proposed by Pawlak [17], is used in various fields. The keywords that characterize methods of rough sets are indiscernibility and discernibility of objects, equivalence classes, and rough approximations expressed by lower and upper ones. The classical framework of rough sets is based on information tables containing only deterministic information, but not containing imprecise information. As a matter of fact, real tables usually contain imprecise information [16]. Therefore, lots of research has been made for information tables with imprecise information [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,18,19,20,21,22].

Among them methods of possible worlds are used for dealing with non-deterministic information [15,18,19,20]. The non-deterministic information is expressed by for example, $\{a, b, c\}$. $\{a, b, c\}$ is an or-set; i.e., its value cannot be determined, but is one among a, b, and c.

In this paper, we extend a method of possible equivalence class, described in [14], dealing with a missing value meaning "do not care" in order to handle non-deterministic information. Non-deterministic values contain the missing value as a special case. Therefore, this extension is a generalization of the method of

P. Wen et al. (Eds.): RSKT 2009, LNCS 5589, pp. 76–85, 2009.

possible equivalence classes. Furthermore, we introduce maximal possible equivalence classes in order to effectively calculate rough approximations.

2 Rough Sets under Deterministic Information

A data set is represented as a table, called an information table, where each row represents an object and each column does an attribute. The information table is pair (U, AT), where U is a non-empty finite set of objects called the universe and AT is a non-empty finite set of attributes such that $\forall a \in AT \; : \; U \rightarrow Dom(a)$ where set $Dom(a)$ is the domain of attribute a. In an information table consisting of set AT of attributes, binary relation $IND(\Psi_A)$ for indiscernibility of objects in subset $\Psi \subseteq U$ on subset $A \subseteq AT$ of attributes is:

$$IND(\Psi_A) = \{(o, o') \in \Psi \times \Psi \mid \forall a \in A \;\; a(o) = a(o')\}. \tag{1}$$

This relation, called an indiscernibility relation, is reflexive, symmetric, and transitive. Obviously, $IND(\Psi_A)$ is an equivalence relation. From the indiscernibility relation, equivalence class $E(\Psi_A)_o$ $(= \{o' \mid (o, o') \in IND(\Psi_A)\})$ containing object o is obtained. This is also the set of objects that is indiscernible with object o, called the indiscernible class for object o. Finally, family $\Psi/IND(\Psi_A)$ $(= \{E(\Psi_A)_o \mid o \in \Psi\})$ of equivalence classes on A is derived from indiscernibility relation $IND(\Psi_A)$. This is the classification induced by A.

Using equivalence classes, lower approximation $\underline{Apr}(\Phi_B, \Psi_A)$ and upper approximation $\overline{Apr}(\Phi_B, \Psi_A)$ of $\Phi/IND(\Phi_B)$ by $\Psi/IND(\Psi_A)$[1] are:

$$\underline{Apr}(\Phi_B, \Psi_A) = \{e \mid \exists e' \; e \subseteq e' \wedge e \in \Psi/IND(\Psi_A) \wedge e' \in \Phi/IND(\Phi_B)\}, \tag{2}$$

$$\overline{Apr}(\Phi_B, \Psi_A) = \{e \mid \exists e' \; e \cap e' \neq \emptyset \wedge e \in \Psi/IND(\Psi_A) \wedge e' \in \Phi/IND(\Phi_B)\}. \tag{3}$$

Expressions in terms of a set of objects are:

$$\underline{apr}(\Phi_B, \Psi_A) = \cup_{e \in \underline{Apr}(\Phi_B, \Psi_A)} e, \quad \overline{apr}(\Phi_B, \Psi_A) = \cup_{e \in \overline{Apr}(\Phi_B, \Psi_A)} e. \tag{4}$$

where apr is used for the expressions by a set of objects while Apr by a family of equivalence classes.

3 Methods of Possible Worlds

In methods of possible worlds, the established ways addressed in the previous section are applied to each possible table, and then the results from the possible tables are aggregated. It is a possible table that every non-deterministic value is replaced by an element comprising the value. When non-deterministic values are contained on set A of attributes in information table T, set $rep(T)_A$ of possible tables on A is:

[1] U_A and U_B are used in place of Ψ_A and Φ_B when sets Ψ and Φ of objects are equal to U, respectively.

$$rep(T)_A = \{pt_1, \ldots, pt_n\}, \tag{5}$$

where each possible table pt_i has an equal possibility that it is the actual one, n is equal to $\Pi_{i=1,m} l_i$, the number of non-deterministic values is m, and each of them has $l_i (i = 1, m)$ elements.

All possible tables consist of deterministic values on A. Family $U/IND(U_A)_{pt_i}$ of equivalence classes on A is obtained from each possible table pt_i. Family $U/IND(U_A)$ of equivalence classes in T is the union of $U/IND(U_A)_{pt_i}$:

$$U/IND(U_A) = \cup_i U/IND(U_A)_{pt_i}. \tag{6}$$

To obtain lower and upper approximations, the classical method addressed in the previous section is applied to possible tables. Let $\underline{Apr}(U_B, U_A)_{pt_i}$ and $\overline{Apr}(U_B, U_A)_{pt_i}$ denote lower and upper approximations of $U/IND(U_B)_{pt_i}$ by $U/IND(U_A)_{pt_i}$ in possible table pt_i. Lower approximation $\underline{Apr}\ (U_B, U_A)$ and upper approximation $\overline{Apr}(U_B, U_A)$ in information table T are the unions of $\underline{Apr}(U_B, U_A)_{pt_i}$ and $\overline{Apr}(U_B, U_A)_{pt_i}$, respectively:

$$\underline{Apr}(U_B, U_A) = \cup_i \underline{Apr}(U_B, U_A)_{pt_i}, \quad \overline{Apr}(U_B, U_A) = \cup_i \overline{Apr}(U_B, U_A)_{pt_i}. \tag{7}$$

We adopt the results by the methods of possible worlds as a correctness criterion for extended methods; namely, $q(T) = \cup q'(rep(T)_A)$, where q' is a classical method and q is an extended method of q' that is described in Section 2.

4 Extending Method of Possible Equivalence Classes

We extend methods of possible equivalence classes described in [14] to deal with non-deterministic information. To handle indiscernibility and discernibility for non-deterministic values, we divide universe U into two sets U_a^d and U_a^{nd} on attribute a. U_a^d and U_a^{nd} consists of objects whose value of attribute $a \in A$ is deterministic and non-deterministic, respectively. For set U_a^d, we obtain family $U_a^d/IND(U_a^d)$ of equivalence classes on attribute a by using the classical method addressed in Section 2. Family $Poss(U/IND(U_A))$ of possible equivalence classes on attribute A is:

$$Poss(U/IND(U_A)) = \cap_{a \in A} \{e \mid e \in Poss(U/IND(U_a))\} \backslash \{\emptyset\}, \tag{8}$$

$$Poss(U/IND(U_a)) = \{e \cup e' \mid e \in U_a^d/IND(U_a^d) \wedge o \in e \wedge e' \in PE^{nd}_{a(o),a}\} \cup_v PE^{nd}_{a=v} \backslash \{\emptyset\}, \tag{9}$$

where $PE^{nd}_{a(o),a}$ and $PE^{nd}_{a=v}$ are the power sets of $E^{nd}_{a(o),a}$ and $E^{nd}_{a=v}$, respectively. $E^{nd}_{a(o),a}$ is the set of objects that have a non-deterministic value that may be equal to deterministic value $a(o)$ on attribute a. $E^{nd}_{a=v}$ is the set of objects that have a non-deterministic value that may be equal to deterministic value v that is an element of non-deterministic values, but does not belong to set V_a^d of deterministic values on attribute a.

$$E^{nd}_{a(o),a} = \{o' \mid o' \in U^{nd}_a \wedge o \in U^d_a \wedge a(o) \in a(o')\}, \tag{10}$$

$$E^{nd}_{a=v} = \{o \mid o \in U^{nd}_a \wedge v \in o(a) \wedge v \in (V^{nd}_a - V^d_a)\}, \tag{11}$$

where $V^d_a = \{o(a) \mid o \in U^d_a\}$ and $V^{nd}_a = \{e \mid e \in o(a) \wedge o \in U^{nd}_a\}$; namely, V^{nd}_a is the set of elements that comprise the non-deterministic values on attribute a. When all non-deterministic values are missing values on attribute a, $E^{nd}_{a(o),a}$ reduces to the set of objects having a missing value and also $E^{nd}_{a=v}$ does under $|V^d_a| < |Dom(a)|$ for cardinality. So, formulae (8) and (9) reduces to the formulae for missing values addressed in [14].

We can express indiscernibility and discernibility of non-deterministic values by using the possible equivalence classes. This is because an object $o \in U^{nd}_a$ is indiscernible with the other objects in a possible equivalence class $e \in Poss(U/IND(U_a))$ on a if o is included in e, otherwise it is discernible.

Using families of possible equivalence classes, we obtain lower and upper approximations $\underline{Apr}(U_B, U_A)$ and $\overline{Apr}(U_B, U_A)$ of $Poss(U/IND(U_B))$ by $Poss(U/IND(U_A))$:

$$\underline{Apr}(U_B, U_A) = \{e \mid \exists e' \; e \subseteq e' \wedge e \in Poss(U/IND(U_A)) \wedge e' \in Poss(U/IND(U_B))\}, \tag{12}$$

$$\overline{Apr}(U_B, U_A) = \{e \mid \exists e' \; e \cap e' \neq \emptyset \wedge e \in Poss(U/IND(U_A)) \wedge e' \in Poss(U/IND(U_B))\}. \tag{13}$$

For expressions in terms of a set of objects, the same expressions as in Section 2 are used.

Proposition 1

The lower and upper approximations that are obtained by the method of possible equivalence classes coincide with ones obtained by the method of possible worlds.

Proof

The proof is similar to that of Proposition 6 in [12].

Example 1

Let information table T be obtained as follows:

T

O	a_1	a_2	a_3	a_4
1	x	u	1	a
2	x	u	1	a
3	x	u	1	a
4	x	$\{u, v\}$	1	a
5	$\{x, y\}$	$\{v, w\}$	2	b

pt_1

O	a_1	a_2	a_3	a_4
1	x	u	1	a
2	x	u	1	a
3	x	u	1	a
4	x	$\{u, v\}$	1	a
5	x	$\{v, w\}$	2	b

pt_2

O	a_1	a_2	a_3	a_4
1	x	u	1	a
2	x	u	1	a
3	x	u	1	a
4	x	$\{u, v\}$	1	a
5	y	$\{v, w\}$	2	b

In information table T, $U = \{o_1, o_2, o_3, o_4, o_5\}$. Let domains $Dom(a_1)$, $Dom(a_2)$, $Dom(a_3)$ and $Dom(a_4)$ of attributes a_1, a_2, a_3 and a_4 be $\{x, y\}$, $\{u, v, w\}$, $\{1, 2\}$ and $\{a, b\}$, respectively. We obtain two possible tables pt_1 and pt_2 from T on a_1, because non-deterministic value $\{x, y\}$ on attribute a_1 of object o_5 is replaced by x or y. o_5 is indiscernible with the other objects on a_1 in pt_1, whereas o_5 is discernible in pt_2. In other words, pt_1 corresponds to the case where o_5 is indiscernible with the other objects on a_1, whereas pt_2 is to the case where o_5 is discernible with the other objects. Families of equivalence classes on attribute a_1 in possible tables pt_1 and pt_2 are:

$$U/IND(U_{a_1})_{pt_1} = \{o_1, o_2, o_3, o_4, o_5\}, \;\; U/IND(U_{a_1})_{pt_2} = \{\{o_1, o_2, o_3, o_4\}, \{o_5\}\}.$$

Let Φ be $\{o_1, o_2, o_3, o_4\}$ for simplicity.

$$\underline{Apr}(\Phi, U_{a_1})_{pt_1} = \{\emptyset\}, \;\; \overline{Apr}(\Phi, U_{a_1})_{pt_1} = \{\{o_1, o_2, o_3, o_4, o_5\}\}.$$
$$\underline{Apr}(\Phi, U_{a_1})_{pt_2} = \{\{o_1, o_2, o_3, o_4\}\}, \;\; \overline{Apr}(\Phi, U_{a_1})_{pt_2} = \{\{o_1, o_2, o_3, o_4\}\}.$$

Using (6),

$$\underline{Apr}(\Phi, U_{a_1}) = \{\{o_1, o_2, o_3, o_4\}\}, \overline{Apr}(\Phi, U_{a_1}) = \{\{o_1, o_2, o_3, o_4\}, \{o_1, o_2, o_3, o_4, o_5\}\},$$

Using (4),

$$\underline{apr}(\Phi, U_{a_1}) = \{o_1, o_2, o_3, o_4\}. \;\; \overline{apr}(\Phi, U_{a_1}) = \{o_1, o_2, o_3, o_4, o_5\}.$$

For attribute a_1 in information table T,

$$U^d_{a_1} = \{o_1, o_2, o_3, o_4\}, \; U^d_{a_1}/IND(U^d_{a_1}) = \{\{o_1, o_2, o_3, o_4\}\}, \;\; U^{nd}_{a_1} = \{o_5\},$$
$$V^d_{a_1} = \{x\}, \;\; E^{nd}_{x,a_1} = \{o_5\}, \;\; V^{nd}_{a_1} = \{x, y\}, \;\; V^{nd}_{a_1} - V^d_{a_1} = \{y\}, \;\; E^{nd}_{a_1=y} = \{o_5\}.$$

Power sets PE^{nd}_{x,a_1} and $PE^{nd}_{a_1=y}$ of E^{nd}_{x,a_1} and $E^{nd}_{a_1=y}$ are $\{\emptyset, \{o_5\}\}$. By using formula (9), the family of possible equivalence classes on attribute a_1 is:

$$Poss(U/IND(U_{a_1})) = \{\{o_5\}, \{o_1, o_2, o_3, o_4\}, \{o_1, o_2, o_3, o_4, o_5\}\}.$$

Using (12) and (13), rough approximations of Φ are.

$$\underline{Apr}(\Phi, U_{a_1}) = \{\{o_1, o_2, o_3, o_4\}\}, \overline{Apr}(\Phi, U_{a_1}) = \{\{o_1, o_2, o_3, o_4\}, \{o_1, o_2, o_3, o_4, o_5\}\}.$$

Indeed, the lower and upper approximations coincide with ones obtained from the method of possible worlds.

5 Maximal Possible Equivalence Classes

As the number of non-deterministic values increases, the number of possible equivalence classes does exponentially. So, the computational time of the method of possible equivalence classes contains an exponential factor of the number of non-deterministic values. To solve this difficulty, we introduce maximal possible equivalence classes. The family $Poss(U/IND(U_A))_{max}$ of maximal possible equivalence classes on A is:

$$Poss(U/IND(U_A))_{max} = \cap_{a \in A}\{e \mid e \in Poss(U/IND(U_a))_{max}\} \backslash \{\emptyset\}, \quad (14)$$
$$Poss(U/IND(U_a))_{max} =$$
$$\{e \cup e' \mid e \in U^d_a/IND(U^d_a) \wedge o \in e \wedge e' = E^{nd}_{a(o),a}\} \cup_v \{E^{nd}_{a=v}\}. \quad (15)$$

Proposition 2
$Poss(U/IND(U_A))_{max} \subseteq Poss(U/IND(U_A))$.

Proof
If $e \in Poss(U/IND(U_A))_{max}$, $e = \cap_{a \in A} e_a$ and $e_a \in Poss(U/IND(U_a))_{max}$ for all $a \in A$ from (14). IF $e_a \in Poss(U/IND(U_a))_{max}$, $e_a \in Poss(U/IND(U_a))$ from (9) and (15). Thus, this proposition holds.

Proposition 3

$$\text{If } e \in Poss(U/IND(U_A)),\ \ \exists e'\ e \subseteq e' \wedge e' \in Poss(U/IND(U_A))_{max}.$$

Proof
If $e \in Poss(U/IND(U_a))$, $e = e' \cup e''$ or $e \in \cup_v PE^{nd}_{a=v}$ from (9) where $e' \in U^d_a/IND(U^d_a)$ and $e'' \in PE^{nd}_{a(o),a}$. If $e'' \in PE^{nd}_{a(o),a}$, $e'' \subseteq E^{nd}_{a(o),a}$. $(e' \cup E^{nd}_{a(o),a}) \in Poss(U/IND(U_a))_{max}$ from (15). If $e \in \cup_v PE^{nd}_{a=v}$, $\exists e'\ e' \subseteq e'' \wedge e'' \in \cup_v \{E^{nd}_{a=v}\}$. Thus, for all $a \in A$, if $e \in Poss(U/IND(U_a))$, $\exists e'\ e \subseteq e' \wedge e' \in Poss(U/IND(U_a))_{max}$. From (8) and (14), this proposition holds.

Using maximal possible equivalence classes, the lower approximation is:

$$\underline{Apr}(U_B, U_A)_{max} = \cap_{a \in A} \{e \mid e \in \underline{Apr}(U_B, U_a)_{max}\} \backslash \{\emptyset\}, \quad (16)$$

$$\underline{Apr}(U_B, U_a)_{max} = \{e \cup (e' \cap E^{nd}_{a(o),a}) \mid \exists e'\ e \subseteq e' \wedge o \in e \wedge e \in U^d_a/IND(U^d_a) \wedge e' \in Poss(U/IND(U_B))_{max}\} \cup \{e \cap e' \mid \exists e'\ \ e \cap e' \neq \emptyset \wedge e \in \cup_v \{E^{nd}_{a=v}\} \wedge e' \in Poss(U/IND(U_B))_{max}\}. \quad (17)$$

The upper approximation is:

$$\overline{Apr}(U_B, U_A)_{max} = \cap_{a \in A} \{e \mid e \in \overline{Apr}(U_B, U_a)_{max}\} \backslash \{\emptyset\}, \quad (18)$$

$$\overline{Apr}(U_B, U_a)_{max} = \{e \mid \exists e'\ e \cap e' \neq \emptyset \wedge e \in Poss(U/IND(U_a))_{max} \wedge e' \in Poss(U/IND(U_B))_{max}\}. \quad (19)$$

Proposition 4
$\underline{Apr}(U_B, U_A)_{max} \subseteq \underline{Apr}(U_B, U_A)$ and $\overline{Apr}(U_B, U_A)_{max} \subseteq \overline{Apr}(U_B, U_A)$

Proof
If $e \in \underline{Apr}(U_B, U_A)_{max}$, $e = \cap_a e_a$ and $\forall a \in A\ e_a \in \underline{Apr}(U_B, U_a)_{max}$ from (16). If $e_a \in \underline{Apr}(U_B, U_a)_{max}$, $\exists e'\ e_a \subseteq e' \wedge e' \in Poss(U/IND(U_B))_{max}$ from (17). If $e' \in Poss(U/IND(U_B))_{max}$, $e' \in Poss(U/IND(U_B))$ from Proposition 2. So, $e \in \underline{Apr}(U_B, U_A)$. The proof is similar for the upper approximation.

Proposition 5

$$\text{If } e \in \underline{Apr}(U_B, U_A),\ \exists e' e \subseteq e' \wedge e' \in \underline{Apr}(U_B, U_A)_{max},$$
$$\text{If } e \in \overline{Apr}(U_B, U_A),\ \exists e' e \subseteq e' \wedge e' \in \overline{Apr}(U_B, U_A)_{max}.$$

Proof
If $e \in \underline{Apr}(U_B, U_A)$, $\exists e'\ e \subseteq e' \wedge e' \in Poss(U/IND(U_B))$ from (12). If $e' \in Poss(U/IND(U_B))$, $\exists e''\ e' \subseteq e'' \wedge e'' \in Poss(U/IND(U_B))_{max}$ from Proposition 3. Thus, this proposition holds. The proof is similar for the upper approximation.

Proposition 6 (Monotonicity of the accuracy of approximations)
If $A \subseteq B$, $\underline{Apr}(U_C, U_A)_{max} \subseteq \underline{Apr}(U_C, U_B)_{max}$ and $\overline{Apr}(U_C, U_A)_{max} \supseteq \overline{Apr}(U_C, U_B)_{max}$.

Proof
If $A \subseteq B \wedge e \in Poss(U/IND(U_B))_{max}$, $\exists e'\ e \subseteq e' \wedge e' \in Poss(U/IND(U_A))_{max}$ from (14). Thus, this proposition holds.

For expressions by a set of objects, the following proposition holds:

Proposition 7

$$\underline{apr}(U_B, U_A)_{max} = \underline{apr}(U_B, U_A), \quad \overline{apr}(U_B, U_A)_{max} = \overline{apr}(U_B, U_A).$$

Proof
If $o \in \underline{apr}(U_B, U_A)_{max}$, $\exists e\ o \in e \wedge e \in \underline{Apr}(U_B, U_A)_{max}$. If $e \in \underline{Apr}(U_B, U_A)_{max}$, $e \in \underline{Apr}(U_B, U_A)$ from Proposition 4. Thus, $o \in \underline{apr}(U_B, U_A)$. If $o \in \underline{apr}(U_B, U_A)$, $\exists e\ o \in e \wedge e \in \underline{Apr}(U_B, U_A)$. From Proposition 5, if $e \in \underline{Apr}(U_B, U_A)$, $\exists e' e \subseteq e' \wedge e' \in \underline{Apr}(U_B, U_A)_{max}$. Thus, $o \in \underline{apr}(U_B, U_A)_{max}$. The proof is similar for the upper approximation.

From Propositions 1 and 7, the method using maximal possible equivalence classes gives the same rough approximations as the method of possible worlds for expressions in terms of a set of objects.

Furthermore, it is noticeable that the method of possible equivalence classes is also applicable to the case that not only condition attributes but also decision attributes contain non-deterministic values. Thus, we can use the method of possible equivalence classes to obtain rough approximations between arbitrary sets of attributes containing non-deterministic values in information tables. We show such an example below.

Example 2
Let the following information table T' be given.

T'

O	a_1	a_2	a_3
1	x	1	a
2	x	2	b
3	$\{x, y, z\}$	2	b
4	y	3	$\{c, d\}$
5	$\{y, z\}$	3	c
6	$\{z, w\}$	3	c

On a_1 and a_3 in T',

$$U^{nd}_{a_1} = \{o_3, o_5, o_6\},\ U^{d}_{a_1} = \{o_1, o_2, o_4\},\ U^{nd}_{a_3} = \{o_4\},\ U^{d}_{a_3} = \{o_1, o_2, o_3, o_5, o_6\}.$$

$$U/IND(U^{d}_{a_1}) = \{\{o_4\}, \{o_1, o_2\}\},\ U/IND(U^{d}_{a_3}) = \{\{o_1\}, \{o_2, o_3\}, \{o_5, o_6\}\},$$
$$V^{d}_{a_1} = \{x, y\},\ E^{nd}_{x,a_1} = \{o_3\},\ E^{nd}_{y,a_1} = \{o_3, o_5\},\ V^{nd}_{a_1} = \{x, y, z, w\},$$
$$V^{nd}_{a_1} - V^{d}_{a_1} = \{z, w\},\ E^{nd}_{a_1=z} = \{o_3, o_5, o_6\},\ E^{nd}_{a_1=w} = \{o_6\},$$
$$V^{d}_{a_3} = \{a, b, c\},\ E^{nd}_{a,a_3} = E^{nd}_{b,a_3} = \emptyset,\ E^{nd}_{c,a_3} = \{o_3\},$$
$$V^{nd}_{a_3} = \{c, d\},\ V^{nd}_{a_3} - V^{d}_{a_3} = \{d\},\ E^{nd}_{a_3=d} = \{o_4\}.$$

Families of maximal possible equivalence classes on a_1 and a_3 are:

$$Poss(U/IND(U_{a_1}))_{max} = \{\{o_6\}, \{o_1, o_2, o_3\}, \{o_3, o_4, o_5\}, \{o_3, o_5, o_6\}\}.$$
$$Poss(U/IND(U_{a_3}))_{max} = \{\{o_1\}, \{o_4\}, \{o_2, o_3\}, \{o_4, o_5, o_6\}\}.$$

Rough approximations are:

$$\underline{Apr}(U_{a_3}, U_{a_1})_{max} = \{\{o_3\}, \{o_4\}, \{o_6\}, \{o_5, o_6\}, \{o_4, o_5\}\},$$
$$\overline{Apr}(U_{a_3}, U_{a_1})_{max} = \{\{o_6\}, \{o_1, o_2, o_3\}, \{o_3, o_4, o_5\}, \{o_3, o_5, o_6\}\}.$$

Expressions by a set of objects are:

$$\underline{apr}(U_{a_3}, U_{a_1})_{max} = \{o_3, o_4, o_5, o_6\}.\ \ \overline{apr}(U_{a_3}, U_{a_1})_{max} = \{o_1, o_2, o_3, o_4, o_5, o_6\}.$$

Of course, these coincide with ones obtained from the method of possible worlds.

6 Conclusions

We have extended the method of possible equivalence classes to deal with information tables containing non-deterministic information. The method gives the same results as the method of possible worlds. This comes from that the method considers both discernibility and indiscernibility of non-deterministic values by using possible equivalence classes. If an object with non-deterministic values is contained in a possible equivalence class, the object is indiscernible with the other objects in the possible equivalence class, otherwise the object is discernible.

It is not necessary to handle all possible equivalence classes. We get the same rough approximations in terms of a set of objects as the method of possible worlds by using maximal possible equivalence classes. In the method using maximal possible equivalence classes, the computational time does not contain the exponential factor of the number of non-deterministic values.

Furthermore, the method is free from the condition that non-deterministic values occur to only condition attributes. Thus, we can use the method of possible equivalence classes to obtain rough approximations between arbitrary sets of attributes containing non-deterministic values.

References

1. Greco, S., Matarazzo, B., Slowinski, R.: Handling Missing Values in Rough Set Analysis of Multi-attribute and Multi-criteria Decision Problem. In: Zhong, N., Skowron, A., Ohsuga, S. (eds.) RSFDGrC 1999. LNCS (LNAI), vol. 1711, pp. 146–157. Springer, Heidelberg (1999)
2. Grzymala-Busse, J.W.: MLEM2: A New Algorithm for Rule Induction from Imperfect Data. In: Proceedings of the IPMU 2002, 9th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Annecy, France, pp. 243–250 (2002)
3. Grzymała-Busse, J.W.: Data with missing attribute values: Generalization of indiscernibility relation and rule induction. In: Peters, J.F., Skowron, A., Grzymała-Busse, J.W., Kostek, B.z., Świniarski, R.W., Szczuka, M.S. (eds.) Transactions on Rough Sets I. LNCS, vol. 3100, pp. 78–95. Springer, Heidelberg (2004)
4. Grzymala-Busse, J.W.: Characteristic relations for incomplete data: A generalization of the indiscernibility relation. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets IV. LNCS (LNAI), vol. 3700, pp. 58–68. Springer, Heidelberg (2005)
5. Guan, Y.-Y., Wang, H.-K.: Set-valued Information Systems. Information Sciences 176, 2507–2525 (2006)
6. Kryszkiewicz, M.: Rules in Incomplete Information Systems. Information Sciences 113, 271–292 (1999)
7. Latkowski, R.: On Decomposition for Incomplete Data. Fundamenta Informaticae 54, 1–16 (2003)
8. Leung, Y., Li, D.: Maximum Consistent Techniques for Rule Acquisition in Incomplete Information Systems. Information Sciences 153, 85–106 (2003)
9. Nakata, M., Sakai, H.: Checking whether or not rough-set-based methods to incomplete data satisfy a correctness criterion. In: Torra, V., Narukawa, Y., Miyamoto, S. (eds.) MDAI 2005. LNCS, vol. 3558, pp. 227–239. Springer, Heidelberg (2005)
10. Nakata, N., Sakai, H.: Rough Sets Handling Missing Values Probabilistically Interpreted. In: Ślezak, D., Wang, G., Szczuka, M.S., Düntsch, I., Yao, Y. (eds.) RSFDGrC 2005. LNCS, vol. 3641, pp. 325–334. Springer, Heidelberg (2005)
11. Nakata, N., Sakai, H.: Applying Rough Sets to Data Tables Containing Imprecise Information Under Probabilistic Interpretation. In: Greco, S., Hata, Y., Hirano, S., Inuiguchi, M., Miyamoto, S., Nguyen, H.S., Słowiński, R. (eds.) RSCTC 2006. LNCS, vol. 4259, pp. 213–223. Springer, Heidelberg (2006)
12. Nakata, N., Sakai, H.: Lower and upper approximations in data tables containing possibilistic information. In: Peters, J.F., Skowron, A., Marek, V.W., Orłowska, E., Słowiński, R., Ziarko, W.P. (eds.) Transactions on Rough Sets VII. LNCS, vol. 4400, pp. 170–189. Springer, Heidelberg (2007)
13. Nakata, N., Sakai, H.: Applying rough sets to information tables containing probabilistic values. In: Torra, V., Narukawa, Y., Yoshida, Y. (eds.) MDAI 2007. LNCS, vol. 4617, pp. 282–294. Springer, Heidelberg (2007)
14. Nakata, N., Sakai, H.: Rough Sets Approximations in Data Tables Containing Missing Values. In: Proceedings of FUZZ-IEEE 2008. IEEE Press, Los Alamitos (2008)
15. Orłowska, E., Pawlak, Z.: Representation of Nondeterministic Information. Theoretical Computer Science 29, 313–324 (1984)
16. Parsons, S.: Current Approaches to Handling Imperfect Information in Data and Knowledge Bases. IEEE Transactions on Knowledge and Data Engineering 83, 353–372 (1996)

17. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht (1991)
18. Sakai, H.: Effective Procedures for Handling Possible Equivalence Relation in Non-deterministic Information Systems. Fundamenta Informaticae 48, 343–362 (2001)
19. Sakai, H., Nakata, M.: An Application of Discernibility Functions to Generating Minimal Rules in Non-deterministic Information Systems. Journal of Advanced Computational Intelligence and Intelligent Informatics 10, 695–702 (2006)
20. Sakai, H., Okuma, A.: Basic Algorithms and Tools for Rough Non-deterministic Information Systems. Transactions on Rough Sets 1, 209–231 (2004)
21. Słowiński, R., Stefanowski, J.: Rough Classification in Incomplete Information Systems. Mathematical and Computer Modelling 12(10/11), 1347–1357 (1989)
22. Stefanowski, J., Tsoukiàs, A.: Incomplete Information Tables and Rough Classification. Computational Intelligence 17(3), 545–566 (2001)