An Attempt of Object Reduction in Rough Set Theory

著者	KUDO Yasuo, MURAI Tetsuya
journal or publication title	2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS)
year	2019
URL	http://hdl.handle.net/10258/00010111

doi: info:doi/10.1109/SCIS-ISIS.2018.00017

An Attempt of Object Reduction in Rough Set Theory

Yasuo Kudo
College of Information and Systems
Muroran Institute of Technology
Muroran, Japan
kudo@csse.muroran-it.ac.jp

Tetsuya Murai
Dept. of Information Systems Engineering
Chitose Institute of Science and Technology
Chitose, Japan
t-murai@photon.chitose.ac.jp

***Abstract*—Attribute reduction is a popular topic in rough set theory; however, object reduction is not considered popularly. In this paper, from a viewpoint of computing all relative reducts, we introduce a concept of object reduction that reduces the number of objects as long as possible with keeping the results of attribute reduction in the original decision table.**

***Index Terms*—rough set, object reduction, attribute reduction, discernibility matrix**

I. Introduction

Attribute reduction is one of the most popular research topics in the community of rough set theory. In Pawlak's rough set theory [2], attribute reduction computes minimal subsets of condition attributes in a given decision table that keep the classification ability by the condition attributes. Such minimal subsets of condition attributes are called relative reducts, and a computation method of all relative reducts using the discernibility matrix has been proposed [3].

On the other hand, to the best of our knowledge, reduction of objects in some sense is not a popular topic in rough set theory. In this paper, from a viewpoint of computing all relative reducts, we introduce a concept of object reduction that reduces the number of objects as long as possible with keeping the results of attribute reduction in the original decision table.

II. Rough Sets

In this section, we briefly review Pawlak's rough set theory. Contents of this section is based on [1].

A. Decision tables, indiscernibility relations, and lower approximations

Let U be a finite set of objects, C be a finite set of condition attributes, and $\mathsf{d} \not{in}C$ be a decision attribute. The following structure is called a decision table to represent a table-style dataset as the target of rough set-based data analysis:

$$DT = (U, C \cup \{\mathsf{d}\}). \tag{1}$$

Each attribute $\mathsf{a} \in C \cup \{d\}$ is a function $\mathsf{a} : U \to V\mathsf{a}$, where V_{a} is a finite set of values of the attribute a.

It is well known that equivalence relations defined on the set U provides partitions of U. Each subset $B \subseteq C \cup \{\mathsf{d}\}$ of attributes constructs an equivalence relation on U, called an indiscernibility relation with respect to B, as follows:

$$IND(B) = \{(x, y) \mid \mathsf{a}(x) = \mathsf{a}(y), \forall \mathsf{a} \in B\}. \tag{2}$$

From the indiscernibility relation $IND(B)$ and an object $x \in U$, an equivalence class $[x]_B$ is obtained. As we mentioned, the set of all equivalence classes with respect to the indiscernibility relation $IND(B)$, called the quotient set, $U/IND(B) = \{[x]_B \mid x \in U\}$ is a partition of U Particularly, the quotient set by the indiscernibility relation $IND(\{\mathsf{d}\})$ with respect to the decision attribute d is called the set of decision classes and denoted by $\mathcal{D} = \{D_1, \cdots, D_m\}$.

For considering attribute reduction in the given decision table, we introduce the lower approximation for each decision class D_i $(1 \leq i \leq m)$ by a subset $B \subseteq C$ of condition attributes as follows:

$$\underline{B}(D_i) = \{x \in U \mid [x]_B \subseteq D_i\}. \tag{3}$$

The lower approximation $\underline{B}(D_i)$ is the set of objects that are correctly classified to D_i by using the information of B.

B. Relative reducts

From the viewpoint of classification of objects, minimal subsets of condition attributes for classifying all discernible objects to correct decision classes are convenient. Such minimal subsets of condition attributes are called relative reducts of the given decision table.

To formally define the relative reducts, we introduce the concept of positive region. Let $B \subseteq C$ be a set of condition attributes. The positive region of the partition $\mathcal{D}$ by B is defined by

$$POS_B(\mathcal{D}) = \bigcup_{D_i \in \mathcal{D}} \underline{B}(D_i). \tag{4}$$

The positive region $POS_B(\mathcal{D})$ is the set of objects classified to correct decision classes by checking the attribute values in every attribute in B. Particularly, the set $POS_C(\mathcal{D})$ is the set of all discernible objects in the decision table DT.

Here, we define the relative reducts. A set $A \subseteq C$ is called a relative reduct of the decision table DT if the set A satisfies the following two conditions:

1) $\mathrm{POS}_A(\mathcal{D}) = \mathrm{POS}_C(\mathcal{D})$.
2) $\mathrm{POS}_B(\mathcal{D}) \neq \mathrm{POS}_C(\mathcal{D})$ for any proper subset $B \subset A$.

In general, there are plural relative reducts in a decision table. For a given decision table DT, we denote the set of all relative reducts of DT by $RED(DT)$.

C. Discernibility matrix

Discernibility matrix was firstly introduced by Skowron and Rauszer [3] i to extract all relative reducts from a decision table. Suppose that the set of objects U in the decision table DT has n objects. The discernibility matrix DM of DT is a symmetric $n \times n$ matrix and its element at the i-th row and j-th column in DM is the following set of condition attributes:

$$\delta_{ij} = \begin{cases} \{\mathsf{a} \in C \mid \mathsf{a}(x_i) \neq \mathsf{a}(x_j)\}, & \\ \quad \text{if } \mathsf{d}(x_i) \neq \mathsf{d}(x_j) \text{ and} & \\ \quad \{x_i, x_j\} \cap POS_C(\mathcal{D}) \neq \emptyset, & \\ \emptyset, & \text{otherwise.} \end{cases} \tag{5}$$

The element δ_{ij} means that the objects x_i is discernible from the object x_j by comparing at least one attribute $\mathsf{a} \in \delta_{ij}$.

Let $\delta_{ij} = \{\mathsf{a}_1, \cdots, \mathsf{a}_l\}$ be the element of DT at i-th row and j-column. Contents of each element δ_{ij} is represented by the logical formula as follows:

$$L(\delta_{ij}) : \mathsf{a}_1 \vee \cdots \vee \mathsf{a}_l. \tag{6}$$

By constructing the conjunctive normal form from the logical formulas and transforming the formula to the prime implicant, all relative reducts in the decision table are computed. The problem of extracting all relative reducts from the given decision table is, however, an NP-hard problem [3], which concludes that computation of all relative reducts from a decision table with numerous objects and attributes is intractable.

Example 1: Table I is an example of a decision table with the set of objects $U = \{x_1, \cdots, x_6\}$, the set of condition attributes $C = \{\mathsf{c}_1, \cdots, \mathsf{c}_6\}$, and $\mathsf{d} \notin C$ is the decision attribute.

Table II shows the discernibility matrix of the decision table in Table I. In Table II, we represent only the lower triangular part of the discernibility matrix. The element $\delta_{61} = \{\mathsf{c}_2, \mathsf{c}_5\}$ means that the element x_6 with $\mathsf{d}(x_6) = 3$ is discernible from x_1 with $\mathsf{d}(x_1) = 1$ by comparing either the value of the attribute c_2 or c_5 between x_6 and x_1.

From the discernibility matrix in Table II, for example, a logical formula $L(\delta_{61}) = \mathsf{c}_2 \vee \mathsf{c}_5$ is constructed based on the nonempty element δ_{61}. By connecting all the logical formula, the following conjunctive normal form is obtained:

$$\begin{gathered} (\mathsf{c}_1 \vee \mathsf{c}_2 \vee \mathsf{c}_3 \vee \mathsf{c}_5 \vee \mathsf{c}_6) \wedge (\mathsf{c}_1 \vee \mathsf{c}_2 \vee \mathsf{c}_3 \vee \mathsf{c}_4 \vee \mathsf{c}_5 \vee \mathsf{c}_6) \\ \wedge(\mathsf{c}_2 \vee \mathsf{c}_5) \wedge (\mathsf{c}_2 \vee \mathsf{c}_3 \vee \mathsf{c}_5 \vee \mathsf{c}_6) \wedge (\mathsf{c}_3 \vee \mathsf{c}_4 \vee \mathsf{c}_5 \vee \mathsf{c}_6) \\ \wedge(\mathsf{c}_1 \vee \mathsf{c}_5) \wedge (\mathsf{c}_2 \vee \mathsf{c}_3 \vee \mathsf{c}_5 \vee \mathsf{c}_6) \wedge (\mathsf{c}_1 \vee \mathsf{c}_2 \vee \mathsf{c}_3 \vee \mathsf{c}_6) \\ \wedge(\mathsf{c}_3 \vee \mathsf{c}_4 \vee \mathsf{c}_5 \vee \mathsf{c}_6) \wedge (\mathsf{c}_1 \vee \mathsf{c}_3 \vee \mathsf{c}_4 \vee \mathsf{c}_6) \wedge (\mathsf{c}_1 \vee \mathsf{c}_3 \vee \mathsf{c}_5). \end{gathered}$$

This conjunctive normal form has many redundant terms. Such redundant terms are eliminated by using idempotent law $P \wedge P = P$ and absorption law $P \wedge (P \vee Q) = P$, where P and Q are any logical formulas. We then obtain the following simplified conjunctive normal form:

$$\begin{gathered} (\mathsf{c}_2 \vee \mathsf{c}_5) \wedge (\mathsf{c}_3 \vee \mathsf{c}_4 \vee \mathsf{c}_5 \vee \mathsf{c}_6) \\ \wedge(\mathsf{c}_1 \vee \mathsf{c}_5) \wedge (\mathsf{c}_1 \vee \mathsf{c}_2 \vee \mathsf{c}_3 \vee \mathsf{c}_6) \\ \wedge(\mathsf{c}_1 \vee \mathsf{c}_3 \vee \mathsf{c}_4 \vee \mathsf{c}_6). \end{gathered}$$

TABLE I

AN EXAMPLE OF A DECISION TABLE

	c_1	c_2	c_3	c_4	c_5	c_6	d
x_1	1	0	0	0	0	1	1
x_2	0	1	0	0	0	1	1
x_3	0	2	1	0	1	0	2
x_4	0	1	1	1	1	0	2
x_5	0	1	2	0	0	1	1
x_6	1	1	0	0	1	1	3

After repeating the application of distributive law, idempotent law, and absorption law, the following prime implicant is obtained:

$$\begin{gathered} (\mathsf{c}_5 \wedge \mathsf{c}_6) \vee (\mathsf{c}_3 \wedge \mathsf{c}_5) \vee (\mathsf{c}_1 \wedge \mathsf{c}_5) \vee (\mathsf{c}_1 \wedge \mathsf{c}_2 \wedge \mathsf{c}_3) \\ \vee(\mathsf{c}_1 \wedge \mathsf{c}_2 \wedge \mathsf{c}_4) \vee (\mathsf{c}_1 \wedge \mathsf{c}_2 \wedge \mathsf{c}_6) \vee (\mathsf{c}_2 \wedge \mathsf{c}_4 \wedge \mathsf{c}_5). \end{gathered}$$

It concludes that there are seven relative reducts in the decision table, i.e., $\{\mathsf{c}_5, \mathsf{c}_6\}$, $\{\mathsf{c}_3, \mathsf{c}_5\}$, $\{\mathsf{c}_1, \mathsf{c}_5\}$, $\{\mathsf{c}_1, \mathsf{c}_2, \mathsf{c}_3\}$, $\{\mathsf{c}_1, \mathsf{c}_2, \mathsf{c}_4\}$, $\{\mathsf{c}_1, \mathsf{c}_2, \mathsf{c}_6\}$, and $\{\mathsf{c}_2, \mathsf{c}_4, \mathsf{c}_5\}$.

III. PROPOSAL OF OBJECT REDUCTION

In this section, we propose a concept of object reduction in rough set theory from a viewpoint of computing all relative reducts in a given decision table.

A. Definition of object reduction

As we illustrated in Example 1, many elements in the discernibility matrix that corresponds to the logical formulas $L(\delta)$ are redundant for computing relative reducts, and these logical formulas are eliminated by using idempotent law and absorption law. This fact means that many elements in the discernibility matrix may not work for computation of all relative reducts of the given decision table. This fact also indicates that some objects in the decision table may not work for computation of all relative reducts, and from a viewpoint of attribute reduction, such objects may be reducible from the decision table.

In this paper, we introduce a concept of object reduction of the given decision table that keeps all relative reducts identical to the original decision table.

Definition 1: Let $DT = (U, C \cup \{\mathsf{d}\})$ be a given decision table and $RED(DT)$ be the set of all relative reducts of DT. Suppose that $x \in U$ is an object of DT, $DT' = (U \setminus \{x\}, C \cup \{\mathsf{d}\})$ be the decision table that the object x is removed from U, where the domain of each attribute $\mathsf{a} \in C \cup \{\mathsf{d}\}$ is restricted to $U \setminus \{x\}$, i.e., $\mathsf{a} : U \setminus \{x\} \to V_{\mathsf{a}}$, and $RED(DT')$ is the set of all relative reducts of the decision table DT'. The object $x \in U$ is called an irreducible object of DT if and only if $RED(DT') \neq RED(DT)$ holds. The object $x \in U$ that is not an irreducible object of DT is called a possibly reducible object of DT.

By this definition, rejection of an irreducible object from the original decision table DT causes some change of results of attribute reduction from DT. In general, removing an irreducible object $x_i \in U$ in DT makes other objects $x_j \in U$ such that there is no need to discern x_i from x_j, which indicates that the condition attributes in the element δ_{ij} in the

TABLE II
THE DISCERNIBILITY MATRIX OF TABLE I

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	$\emptyset$					
x_2	$\emptyset$	$\emptyset$				
x_3	$\{c_1, c_2, c_3, c_5, c_6\}$	$\{c_2, c_3, c_5, c_6\}$	$\emptyset$			
x_4	$\{c_1, c_2, c_3, c_4, c_5, c_6\}$	$\{c_3, c_4, c_5, c_6\}$	$\emptyset$	$\emptyset$		
x_5	$\emptyset$	$\emptyset$	$\{c_2, c_3, c_5, c_6\}$	$\{c_3, c_4, c_5, c_6\}$	$\emptyset$	
x_6	$\{c_2, c_5\}$	$\{c_1, c_5\}$	$\{c_1, c_2, c_3, c_6\}$	$\{c_1, c_3, c_4, c_6\}$	$\{c_1, c_3, c_5\}$	$\emptyset$

original discernibility matrix DM may not appear in relative reducts.

On the other hand, rejection of a possibly reducible object does not affect to attribute reduction. The term "possibly reducible" means that not all possible reducible objects may be reducible from the original decision table, i.e., rejection of a possibly reducible object in the original decision table DT may make some other possibly reducible object in DT an irreducible object in the resulted decision table DT'.

Example 2: Here, we consider two examples of decision tables based on the original decision table by Table I. The first example is a decision table by removing the object $x_5 \in U$ from Table I and we denote this decision table as DTx_5. To consider computing all relative reducts from the decision table DT_{x_5}, we ignore the row and column of x_5 in the discernibility matrix of DT in Table II, applying idempotent law and absorption law to the logical formula constructed from the discernibility matrix, and the following conjunctive normal form is obtained:

$$(c_2 \vee c_5) \wedge (c_3 \vee c_4 \vee c_5 \vee c_6) \\ \wedge (c_1 \vee c_5) \wedge (c_1 \vee c_2 \vee c_3 \vee c_6) \\ \wedge (c_1 \vee c_3 \vee c_4 \vee c_6).$$

This formula is identical to the conjunctive normal form that we illustrated in Example 1, which concludes that all relative reducts obtained from the decision table DT_{x_5} are identical to the relative reducts from the original decision table DT. Therefore, $RED(DT_{x_5}) = RED(DT)$ holds and the object $x_5 \in U$ is a possibly reducible object of DT.

Next example is a decision table by removing the object $x_4 \in U$ from DT and we denote this decision table as DT_{x_4}. Similar to the case of DT_{x_5}, we ignore the row and column of x_4 in Table II, applying idempotent law and absorption law, and the following conjunctive normal form is obtained:

$$(c_2 \vee c_5) \wedge (c_1 \vee c_5) \wedge (c_1 \vee c_2 \vee c_3 \vee c_6).$$

This formula is transformed to the following prime implicant:

$$(c_1 \wedge c_2) \vee (c_1 \wedge c_5) \vee (c_2 \wedge c_5) \vee (c_3 \wedge c_5) \vee (c_5 \wedge c_6),$$

which concludes that there are five relative reduct in DT_{x_4}, i.e., $\{c_1, c_2\}$, $\{c_1, c_5\}$, $\{c_2, c_5\}$, $\{c_3, c_5\}$, and $\{c_5, c_6\}$. Therefore, $RED(DT_{x_4}) \neq RED(DT)$ holds and the object $x_4 \in U$ is an irreducible object of DT.

B. Properties of possibly reducible objects and irreducible objects

Discussion in Example 2 indicates close relationship between the concept of reducibility of objects and discernibility matrix. In this subsection, we consider theoretical connections between possibly reducible objects and discernibility matrix.

Proposition 1: Let $DT = (U, C \cup \{d\})$ be a given decision table and DM be the discernibility matrix of DT. An object $x_i \in U$ is a possibly reducible object of DT if and only if, for every nonempty element δ_{ij} $(1 \leq j \leq |U|)$, there exists an element δ_{kl} in DM that either $i \neq k$ or $j \neq l$ and $\delta_{kl} \subseteq \delta_{ij}$ hold.

This proposition means that, for every nonempty element δ_{ij} that is required to discern the possibly reducible object x_i from another object x_j, the corresponding logical formula $L(\delta_{ij})$ will be eliminated by the logical formula $L(\delta_{kl})$ by using idempotent law or absorption law. Then, information to discern the object x_i from the object x_j is already included in the element δ_{kl} to discern other objects x_k and x_l, and therefore, discerning x_i from x_j has no meaning from the viewpoint of attribute reduction.

Corollary 1: Let DT be a given decision table and DM be the discernibility matrix of DT. An object $x_i \in U$ is an irreducible object of DT if and only if there exists a nonempty element δ_{ik} in DM such that any nonempty element $\delta \neq \emptyset$ in DM is not a proper subset of δ_{ik}, i.e., $\delta \not\subset \delta_{ik}$ holds.

The original definition of irreducible objects and possibly reducible objects are based on the set of all relative reducts $RED(DT)$ of the given decision table DT. As we mentioned in Section II-C, however, computing all relative reducts is actually impossible from the decision table with numerous objects and attributes. Hence, the original concept of reducibility of objects is not computable.

On the other hand, by Proposition 1, now we can compute possibly reducible objects concretely. This computation is based on making the discernibility matrix and comparison of elements in the discernibility matrix by set inclusion relationship. Consequently, there is no need to compute all relative reducts to find possibly reducible objects.

C. An algorithm for object reduction

In this subsection, we introduce an algorithm for computing a result of object reduction by removing possibly reducible objects in a given decision table as many as possible. Relative reducts of the resulted decision table are identical to the ones of the original decision table.

Algorithm 1 outputs a result of object reduction of the given decision table. Steps 3-9 correspond to elimination of redundant elements δ_{ij} in the discernibility matrix DM by using idempotent law and absorption law. Steps 10-14 remove objects x_i as possibly reducible objects because all elements δ_{ij} to discern x_i from another discernible objects x_j are replaced to empty set, i.e., discerning x_i from other objects is redundant for attribute reduction.

Algorithm 1 dtr: decision table reduction algorithm

Input: decision table $DT = (U, C \cup \{\mathsf{d}\})$
Output: result of object reduction $DT' = (U', C \cup \{\mathsf{d}\})$
1: Compute the discernibility matrix DM of DT
2: $U' \leftarrow U$
3: **for all** $\delta_{kl} \in DM$ **do**
4:     **for all** $\delta_{ij} \in DM$ **do**
5:         **if** $(i \neq k$ **or** $j \neq l)$ **and** $\delta_{kl} \subseteq \delta_{ij}$ **then**
6:             $\delta_{ij} \leftarrow \emptyset$
7:         **end if**
8:     **end for**
9: **end for**
10: **for** $i = 1$ to $|U|$ **do**
11:     **if** $\delta_{ij} = \emptyset$ for all $j \in \{1, \cdots, |U|\}$ **then**
12:         $U' \leftarrow U' \setminus \{x_i\}$
13:     **end if**
14: **end for**
15: **return** $DT' = (U', C \cup \{\mathsf{d}\})$.

Example 3: We show an example of object reduction of Table I by Algorithm 1. Comparing elements in DM by Table II and replacing redundant elements δ_{ij} to empty set, as a result, the following nonempty elements are obtained: $\delta_{61} = \{\mathsf{c}_2, \mathsf{c}_5\}$, $\delta_{62} = \{\mathsf{c}_1, \mathsf{c}_5\}$, $\delta_{42} = \{\mathsf{c}_3, \mathsf{c}_4, \mathsf{c}_5, \mathsf{c}_6\}$, $\delta_{63} = \{\mathsf{c}_1, \mathsf{c}_2, \mathsf{c}_3, \mathsf{c}_6\}$, and $\delta_{64} = \{\mathsf{c}_1, \mathsf{c}_3, \mathsf{c}_4, \mathsf{c}_6\}$. Consequently, $\delta_{5j} = \emptyset$ holds for all $j \in \{1, \cdots, 6\}$, the object x_5 is a possibly redundant object and it is removed from U. As we have shown in Example 2, the resulted decision table generates all relative reducts that are identical to the relative reducts in the original decision table.

D. Application to dataset

We used Algorithm 1 to Zoo dataset in UCI Machine Learning Repository [4]. The Zoo dataset consists of 101 animals (objects) and 17 attributes like "hair", "feather", "egg", etc.. We set the attribute "type" as the decision attribute, and remaining 16 attributes as condition attributes. The decision attribute "type" divides 101 animals to 7 classes like "mammalian", "birds", "reptiles", "amphibians", "fishes", "insects", and "others". We confirmed that 33 relative reducts are obtained from the original Zoo dataset.

As a result of applying Algorithm 1 to the Zoo dataset, 86 objects were removed as possibly reducible objects. We also confirmed that 33 relative reducts were obtained from the resulted decision table and all relative reducts were identical to the relative reducts from the original Zoo dataset.

IV. Conclusion

In this paper, we proposed an approach of object reduction in rough set theory. Our approach is based on removing as many objects as possible with keeping the result of extraction of all relative reducts. We introduced the concept of possibly reducible objects as the objects that do not affect the result of computing all relative reducts by removing the objects from the decision table, and the concept of irreducible objects as the objects that removing the objects changes the result of computing all relative reducts. We showed theoretical connection between possibly reducible objects and the discernibility matrix of the decision table, and also proposed an algorithm to compute object reduction based on the theoretical connection. Experiment result indicates that the proposed approach can efficiently reduce the number of objects with keeping the ability of attribute reduction. Future works are more refinement of theoretical connection between possibly reducible objects and discernibility matrix, and further experiments of the proposed approach using various datasets. To determine the remaining element in the discernibility matrix when using the idempotent law and absorption llaw, object-wise search technique is required to select the remaining element to remove as many objects as possible. This refinement is an important future issue.

Acknowledgment

The authors would like to thank the anonymous reviewers for their helpful comments and suggestions.

References

[1] N. Mori, H. Tanaka, K. Inoue (eds.), Rough Sets and Kansei: Knowledge Acquisition and Reasoning from Kansei Data, Kaibundo 2004 (in Japanese).
[2] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, 1991.
[3] A. Skowron and C. M. Rauszer, The discernibility matrix and functions in information systems, *Intelligent Decision Support: Handbook of Application and Advance of the Rough Set Theory*, R. Słowiński (ed.), Kluwer Academic Publishers, pp.331–362, 1992.
[4] UCI Machine Learning Repository, http://archive.ics.uci.edu/ml/index.php