Region Vector Based Attribute Reducts in Decision-Theoretic Rough Sets

Guoshun Huang[1,2(✉)] and Yiyu Yao[2]

[1] School of Science, Foshan University, Foshan 528000, China
fshgs_72@163.com
[2] Department of Computer Science, University of Regina,
Regina, SK S4S 0A2, Canada
yyao@cs.uregina.ca

Abstract. When removing some attributes, the partition induced by a smaller set of attributes will be coarser and the decision regions may be changed. In this paper, we analyze the decision region changes when removing attributes and propose a new type of attribute reducts from the point of view of vector based three-way approximations of a partition. We also present a reduct construction method by using a discernibility matrix.

Keywords: Attribute reduction · Decision-theoretic rough sets · Discernibility matrix

1 Introduction

Decision-theoretic rough set (DTRS) models [15,16] are probabilistic generalizations of Pawlak rough sets [9]. One feature of DTRS models is the tolerance of decision errors. In contrast to Pawlak approximations [9], both probabilistic positive and negative regions are not certain and contain errors [1,8,13]. Consequently, the notion of attribute reducts becomes more complicated. In addition to the condition of positive region preservation required by a Pawlak attribute reduct [10], we must also consider boundary and negative regions. This leads to the notion of region vector based attribute reducts.

Compared with Pawlak's three regions, probabilistic regions of DTRS models are not monotonic with respect to the set inclusion relation on sets of attributes [5,18]. Definitions of, and methods for constructing, Pawlak attribute reducts may not be appropriate for DTRS models. To address the problem of non-monotonicity, many authors proposed different types of attribute reducts, such as minimum cost attribute reducts by Jia et al. [2,4], non-monotonic attribute reducts by Li et al. [5], cost-sensitive attribute reducts by Liao et al. [6], decision region distribution preservation reducts by Ma et al. [7], and so on.

Zhao et al. [18] proposed a general definition of an attribute reduct as a minimal set of condition attributes satisfying properties given in terms of a set

of evaluation measures. Jia et al. [3] studied systematically several specific classes of attribute reducts based on a general definition of attribute reducts. Yao [14] provided a more general conceptual definition of reducts of a set, embracing notions of attribute reducts, attribute-value pair reducts, and rule reducts.

Based on results from these studies, in this paper we study attribute reducts by considering changes of decision regions. By using region vector based three-way approximations, we propose three new definitions of attribute reducts. The new definitions require that an attribute reduct should perverse high confidence rules. A method for constructing an attribute reduct is proposed by using the notion of a discernibility matrix introduced by Skowron and Rauszer [11].

2   Three-Way Approximations of a Classification

An information table is defined by $IT = (U, AT, \{V_a | a \in AT\}, \{I_a | a \in AT\})$, where U is a finite nonempty set of objects, AT is a finite nonempty set of attributes, V_a is a nonempty set of values for $a \in AT$, and $I_a : U \to V_a$ is an information function that maps an object in U to one value in V_a. If $AT = C \cup D$, where C is a finite set of condition attributes, D is a finite set of decision attributes and $C \cap D = \emptyset$, the information table is called a decision table, denoted by DT.

For a subset of attributes $A \subseteq AT$, one defines an equivalence relation on U as follows:

$$xR_A y \Leftrightarrow \forall a \in A(I_a(x) = I_a(y)). \tag{1}$$

The equivalence relation R_A induces a partition of U, denoted by $U/R_A = \{[x]_{R_A} \mid x \in U\}$ and U/A or π_A for short. The equivalence class containing x is given by $[x]_A = \{y \mid \forall a \in A(I_a(x) = I_a(y))\}$. According to Pawlak [10], a partition induced by a subset of attributes is called a classification of U. The classification based on the set of decision attributes is given by U/D or π_D.

For a subset of objects $X \subseteq U$ and a subset of attributes $A \subseteq AT$, Let $\Pr(X|[x]_A)$ denote the conditional probability of an object belonging to X given that the object belongs to $[x]_A$. This probability may be simply estimated as $\Pr(X|[x]_A) = |X \cap [x]_A| / |[x]_A|$, where $|\cdot|$ denotes the cardinality of a set. In terms of conditional probability, the main results of DTRS models are approximations of a set through a pair of lower and upper approximations or three pair-wise disjoint positive, boundary and negative regions. In this paper, we adopt the formulation with three-way approximations [13,14] by using a slightly different notional system.

Definition 1. *In a decision table $DT = (U, AT = C \cup D, \{V_a | a \in AT\}, \{I_a | a \in AT\})$, given a subset of attributes $A \subseteq C$ and a pair of thresholds $0 \leq \beta < \alpha \leq 1$, the following sets:*

$$
\begin{aligned}
\mathrm{POS}_{(\alpha,\beta)}(X|\pi_A) &= \{x \in U \mid \Pr(X|[x]_A) \geq \alpha\}, \\
\mathrm{BND}_{(\alpha,\beta)}(X|\pi_A) &= \{x \in U \mid \beta < \Pr(X|[x]_A) < \alpha\}, \\
\mathrm{NEG}_{(\alpha,\beta)}(X|\pi_A) &= \{x \in U \mid \Pr(X|[x]_A) \leq \beta\},
\end{aligned}
\tag{2}
$$

are called (α, β)-positive, boundary and negative regions of a subset of objects X with respect to the partition π_A.

The three regions are pair-wise disjoint and their union is U. Some of the three regions may be the empty set. Thus, the family of these regions may not necessarily be a partition of U.

One can extend the three-way approximations of a set to three-way approximations of a classification $\pi_D = \{D_1, D_2, \cdots, D_r\}$. A straightforward way is to define three-way approximations component-wise.

Definition 2. *In a decision table DT, given a subset of attributes $A \subseteq C$ and a pair of thresholds $0 \le \beta < \alpha \le 1$, the vector based (α, β)-positive, boundary and negative regions of π_D with respect to the partition π_A are given by:*

$$\overrightarrow{\mathrm{POS}}_{(\alpha,\beta)}(\pi_D|\pi_A) = (\mathrm{POS}_{(\alpha,\beta)}(D_1|\pi_A), \cdots, \mathrm{POS}_{(\alpha,\beta)}(D_r|\pi_A)),$$
$$\overrightarrow{\mathrm{BND}}_{(\alpha,\beta)}(\pi_D|\pi_A) = (\mathrm{BND}_{(\alpha,\beta)}(D_1|\pi_A), \cdots, \mathrm{BND}_{(\alpha,\beta)}(D_r|\pi_A)),$$
$$\overrightarrow{\mathrm{NEG}}_{(\alpha,\beta)}(\pi_D|\pi_A) = (\mathrm{NEG}_{(\alpha,\beta)}(D_1|\pi_A), \cdots, \mathrm{NEG}_{(\alpha,\beta)}(D_r|\pi_A)). \quad (3)$$

By taking a component-wise union of regions in two vectors $\overrightarrow{\mathrm{POS}}_{(\alpha,\beta)}(\pi_D|\pi_A)$ and $\overrightarrow{\mathrm{BND}}_{(\alpha,\beta)}(\pi_D|\pi_A)$, we obtain a non-negative region vector:

$$\overrightarrow{\neg\mathrm{NEG}}_{(\alpha,\beta)}(\pi_D|\pi_A) = (\neg\mathrm{NEG}_{(\alpha,\beta)}(D_1|\pi_A), \cdots, \neg\mathrm{NEG}_{(\alpha,\beta)}(D_r|\pi_A)), \quad (4)$$

where $\neg\mathrm{NEG}_{(\alpha,\beta)}(D_j|\pi_A) = \mathrm{POS}_{(\alpha,\beta)}(D_j|\pi_A) \cup \mathrm{BND}_{(\alpha,\beta)}(D_j|\pi_A)$, $j = 1, \cdots, r$.

Another definition is based on the union of the three-way approximations of all decision classes, we call it a set based definition.

Definition 3. *In a decision table DT, given a subset of attributes $A \subseteq C$ and a pair of thresholds $0 \le \beta < \alpha \le 1$, the set based (α, β)-positive, boundary and negative regions of π_D with respect to the partition π_A are given by:*

$$\mathrm{POS}_{(\alpha,\beta)}(\pi_D|\pi_A) = \bigcup_{j=1}^{r} \mathrm{POS}_{(\alpha,\beta)}(D_j|\pi_A),$$
$$= \{x \in U \mid \exists D_j \in \pi_D(\mathrm{Pr}(D_j|[x]_A) \ge \alpha)\};$$
$$\mathrm{BND}_{(\alpha,\beta)}(\pi_D|\pi_A) = \bigcup_{j=1}^{r} \mathrm{BND}_{(\alpha,\beta)}(D_j|\pi_A),$$
$$= \{x \in U \mid \exists D_j \in \pi_D(\beta < \mathrm{Pr}(D_j|[x]_A) < \alpha)\};$$
$$\mathrm{NEG}_{(\alpha,\beta)}(\pi_D|\pi_A) = U - \mathrm{POS}_{(\alpha,\beta)}(\pi_D|\pi_A) \cup \mathrm{BND}_{(\alpha,\beta)}(\pi_D|\pi_A),$$
$$= \{x \in U \mid \forall D_j \in \pi_D(\mathrm{Pr}(D_j|[x]_A) \le \beta)\}. \quad (5)$$

A set based (α, β)-non-negative region is defined as the union of positive and boundary regions:

$$\neg\mathrm{NEG}_{(\alpha,\beta)}(\pi_D|\pi_A) = \bigcup_{j=1}^{r} \{x \in U | \exists D_j \in \pi_D(\mathrm{Pr}(D_j \mid [x]_A) > \beta)\}. \quad (6)$$

As $\mathrm{POS}_{(\alpha,\beta)}(\pi_D|\pi_A)$ only depends on the threshold α, Li et al. [5] call it the α-positive region. For convenience, we still call it (α,β)-positive region. In the special case when $\alpha = 1$ and $\beta = 0$, $\mathrm{POS}_{(1,0)}(\pi_D|\pi_A)$ and $\mathrm{BND}_{(1,0)}(\pi_D|\pi_A)$ are, in fact, the positive and boundary regions in Pawlak rough set model, respectively. In this case, $\mathrm{NEG}_{(1,0)}(\pi_D|\pi_A) = \emptyset$.

One can construct decision or classification rules from the three regions. In general, a rule in rough set theory can be expressed in the form of $[x]_A \rightarrow D_j$, stating that an object with description $[x]_A$ would be in the decision class D_j. Given a rule $[x]_A \rightarrow D_j$, $\mathrm{Pr}(D_j|[x]_A)$ is its confidence and $\mathrm{Pr}([x]_A|D_j)$ its coverage [17].

Skowron [12] introduced a concept of generalized decisions. For an object x, the generalized decision for x is the set all possible decisions for objects in the equivalence class $[x]_A$. For DTRS, Zhao et al. [18] introduced the notion of (α,β)-positive, boundary and non-negative decisions. Specifically, for an equivalence class $[x]_A \in \pi_A$, we have:

$$D_{\mathrm{POS}(\alpha,\beta)}([x]_A) = \{D_j \in \pi_D \mid (\mathrm{Pr}(D_j|[x]_A) \geq \alpha)\},$$
$$D_{\mathrm{BND}(\alpha,\beta)}([x]_A) = \{D_j \in \pi_D \mid (\beta < \mathrm{Pr}(D_j|[x]_A) < \alpha)\},$$
$$D_{\neg\mathrm{NEG}(\alpha,\beta)}([x]_A) = D_{\mathrm{POS}(\alpha,\beta)}([x]_A) \cup D_{\mathrm{BND}(\alpha,\beta)}([x]_A). \quad (7)$$

They are probabilistic versions of the generalized decisions.

3   Region Vector Based Attribute Reducts for DTRS Models

In existing studies, the set based regions are widely used in defining attribute reducts for DTRS models. The problem of searching for an attribute reduct is to remove redundant attributes. By successively removing attributes, the resulting partitions will become coarser. As a result, the decision regions will be changed. Compared with Pawlak rough set model, the monotonicity of decision region change no longer holds for DTRS models. It is therefore very important to analyze how decision region changes with the decrease of attributes.

3.1   An Analysis of Set Based Decision Regions

For the set based three regions, it is easy to verify the following properties:

(1) $\mathrm{NEG}_{(\alpha,\beta)}(\pi_D|\pi_A)$ may not be empty,

(2) $\mathrm{NEG}_{(\alpha,\beta)}(\pi_D|\pi_A) \cap \mathrm{POS}_{(\alpha,\beta)}(\pi_D|\pi_A) = \emptyset$,

(3) $\mathrm{NEG}_{(\alpha,\beta)}(\pi_D|\pi_A) \cap \mathrm{BND}_{(\alpha,\beta)}(\pi_D|\pi_A) = \emptyset$,

(4) $\mathrm{POS}_{(\alpha,\beta)}(\pi_D|\pi_A)$ and $\mathrm{BND}_{(\alpha,\beta)}(\pi_D|\pi_A)$ may have an overlap,

(5) When $\alpha > 0.5, \forall D_i, D_j \in \pi_D, i \neq j,$
$$\mathrm{POS}_{(\alpha,\beta)}(D_i|\pi_A) \cap \mathrm{POS}_{(\alpha,\beta)}(D_j|\pi_A) = \emptyset,$$

(6) When $\alpha \leq 0.5,$ there may exist decision classes such that
$$\mathrm{POS}_{(\alpha,\beta)}(D_i|\pi_A) \cap \mathrm{POS}_{(\alpha,\beta)}(D_j|\pi_A) \neq \emptyset, \quad i \neq j.$$

These properties are essential to our study of attribute reducts in DTRS models.

When we remove some attributes, the confidence of decision rules, as defined by the conditional probability may change. Consider a subset of condition attributes $A \subseteq C$. By definition, an equivalence class of π_A is the union of a family of equivalence classes of C. That is, $[x]_A = \cup_{y \in [x]_A} [y]_C = \cup_{k=1}^m [y_k]_C$, where $[y_1]_C, \ldots, [y_m]_C$ are the family of distinct equivalence classes of π_C such that their union is $[x]_A$. Based on this result, one can easily verify the next lemma.

Lemma 1. *For $D_j \in \pi_D$, we have:*

(1) $\Pr(D_j|[x]_A) = \sum_{k=1}^m \dfrac{|[y_k]_C|}{|[x]_A|} \Pr(D_j|[y_j]_C),$

(2) $\min\{\Pr(D_j|[y]_C) \mid y \in [x]_A\} \leq \Pr(D_j|[x]_A) \leq \max\{\Pr(D_j|[y]_C) \mid y \in [x]_A\}.$

By removing some attributes, we in fact combine a family rules into a single rule. According to Lemma 1, the confidence of the new rule $[x]_A \to D_j$ is the weighted sum of the confidence of individual rules $[y_k]_C \to D_j$, $k = 1, \ldots, m$.

By the set based definition of three regions, $[x]_C$ is put into the (α, β)-positive region $\mathrm{POS}_{(\alpha,\beta)}(\pi_D|\pi_C)$ if $\Pr(D_j|[x]_C) \geq \alpha$ for some $D_j \in \pi_D$. Equivalence classes in positive regions induced by π_C with different decision set may merge into a new equivalence class π_A when removing attributes and the new equivalence class may be in the negative region $\mathrm{NEG}_{(\alpha,\beta)}(\pi_D|\pi_C)$, as shown by the following example from [2].

Example 1. Consider a decision Table 1, where $U = \{x_1, \cdots, x_9\}$, $C = \{c_1, c_2\}$, and $D = \{d\}$. The partition induced by the set of condition attributes is given by $U/C = \{C_1, C_2, C_3, C_4\}$, where $C_1 = \{x_1, x_4\}$, $C_2 = \{x_2, x_6, x_8\}$, $C_3 = \{x_3\}$, and $C_4 = \{x_5, x_7, x_9\}$. The decision partition is given by $U/D = \{D_1, D_2, D_3\}$, where $D_1 = \{x_1, x_2\}$, $D_2 = \{x_3, x_4, x_5\}$, and $D_3 = \{x_6, x_7, x_8, x_9\}$. Let $A = \{c_1\}$, then $U/A = \{A_1, A_2\}$, where $A_1 = C_1 \cup C_2$ and $A_2 = C_3 \cup C_4$.

Assume that $\alpha = 0.6$ and $\beta = 0.5$. We have $\mathrm{POS}_{(0.6,0.5)}(\pi_D|\pi_C) = C_2 \cup C_3 \cup C_4$ and $\mathrm{NEG}_{(0.6,0.5)}(\pi_D|\pi_A) = U$. The equivalence classes C_3 and C_4 in partition U/C is merged into A_2 in partition U/A when we remove attribute c_2. The equivalence class A_2 is put into negative region $\mathrm{NEG}_{(0.6,0.5)}(\pi_D|\pi_A)$.

When $\alpha = 0.6$ and $\beta = 0.4$, the equivalence class A_2 is put into the boundary region $\mathrm{BND}_{(0.6,0.4)}(\pi_D|\pi_A)$. When $\alpha = 0.5$ and $\beta = 0.3$, A_2 is put into positive region $\mathrm{POS}_{(0.5,0.3)}(\pi_D|\pi_A)$.

Analogously, some equivalence classes in the positive or negative regions merge into a new equivalence class, which may be put into the positive, boundary or negative region. There is one exception, that is, the merge of equivalence classes in the negative region is still in the negative region according to the definition of negative region. Table 2 summarizes the main results, where the second column gives the regions of equivalences before the merge, and the third column gives the regions of the new equivalence class.

Table 1. A decision table

U	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
c_1	1	1	0	1	0	1	0	1	0
c_2	1	0	1	1	0	0	0	0	0
d	0	0	1	1	1	2	2	2	2

Table 2. Changes of decision regions after removing attributes

Cases	The original regions	The new regions after removing attributes
(1)	All positive	Positive, boundary or negative
(2)	Positive and boundary	Positive, boundary or negative
(3)	Positive and negative	Positive, boundary or negative
(4)	All boundary	Positive, boundary or negative
(5)	Boundary and negative	Positive, boundary or negative
(6)	All negative	negative
(7)	Positive,boundary and negative	Positive, boundary or negative

3.2   Region Vector Based Attribute Reducts

According to the previous analysis, a set based region may become larger, unchanged or smaller when we remove some condition attributes. This results in difficulties in constructing an attribute reduct for DTRS models. For example, when equivalence classes in positive and negative regions merge, the new equivalence class will be put into positive, boundary or negative regions. If it is put into the positive region, the new positive region will be larger. Such a move may not be desirable because it puts original lower confidence rules into positive region. On the other hand, if the merged equivalence class is put into the boundary or negative regions, it cannot guarantee that a positive rule is still a positive rule after removing attributes.

When equivalence classes with the same decision set are combined, the new equivalence class can put into the positive region according to Lemma 1. An original positive rule is still a positive rule after reducing attributes. The generality stay the same and there is no risk to misclassify a low confidence rule as a high confidence rule.

Similarly, if equivalence classes in boundary regions with the same decision set are merged, the merged equivalence class is still in the boundary region by Lemma 1. This will not change the positive, boundary and negative regions. When equivalence classes in negative region are merged, the merged equivalence classes must be put into negative region. This does not change the generality and nor increases the cost of decision risk.

Based on the above analysis, we propose a new type of attribute reducts based on a vector representation of three regions.

Definition 4. *In a decision table DT, given a subset of attributes $R \subseteq C$ and a pair of thresholds $0 \le \beta < \alpha \le 1$, R is an (α, β)-vregion attribute reduct of C with respect to the set of decision attributes D if it satisfies the following two conditions:*

(1) $\overrightarrow{\text{vregion}}_{(\alpha,\beta)}(\pi_D|\pi_R) = \overrightarrow{\text{vregion}}_{(\alpha,\beta)}(\pi_D|\pi_C);$

(2) *for any* $R' \subset R, \overrightarrow{\text{vregion}}_{(\alpha,\beta)}(\pi_D|\pi_{R'}) \ne \overrightarrow{\text{vregion}}_{(\alpha,\beta)}(\pi_D|\pi_C).$

Especially, if vregion is a positive, boundary, negative or non-negative region, then R is called an (α, β)-positive, boundary, negative or non-negative region attribute reduct of C, respectively. If R only satisfies condition (1), it is called an (α, β)-positive, boundary, negative or non-negative region consistent attribute subset of C.

Theorem 1. $\overrightarrow{\text{POS}}_{(1,0)}(\pi_D|\pi_R) = \overrightarrow{\text{POS}}_{(1,0)}(\pi_D|\pi_C)$ *if and only if* $\text{POS}_{(1,0)}(\pi_D|\pi_R) = \text{POS}_{(1,0)}(\pi_D|\pi_C)$.

Proof. ($\Rightarrow$) If $\overrightarrow{\text{POS}}_{(1,0)}(\pi_D|\pi_R) = \overrightarrow{\text{POS}}_{(1,0)}(\pi_D|\pi_C)$, by definitions of these regions, we immediately have $\text{POS}_{(1,0)}(\pi_D|\pi_R) = \text{POS}_{(1,0)}(\pi_D|\pi_C)$.
($\Leftarrow$) By definition, $\text{POS}_{(1,0)}(D_i|\pi_R) \cap \text{POS}_{(1,0)}(D_j|\pi_R) = \emptyset$ for any $D_i, D_j \in \pi_D, i \ne j$ and $\text{POS}_{(1,0)}(D_j|\pi_R) \subseteq \text{POS}_{(1,0)}(D_j|\pi_C)$. Hence, if $\text{POS}_{(1,0)}(\pi_D|\pi_R) = \text{POS}_{(1,0)}(\pi_D|\pi_C)$, we have $\text{POS}_{(1,0)}(D_j|\pi_R) = \text{POS}_{(1,0)}(D_j|\pi_C)$. It follows that $\overrightarrow{\text{POS}}_{(1,0)}(\pi_D|\pi_R) = \overrightarrow{\text{POS}}_{(1,0)}(\pi_D|\pi_C)$.

Theorem 1 shows that, in Pawlak rough set model, the two representation forms are equivalent in judging whether R is a positive region preservation consistent set of C. For boundary region, we can only get a one-way inference. That is, $\overrightarrow{\text{BND}}_{(\alpha,\beta)}(\pi_D|\pi_R) = \overrightarrow{\text{BND}}_{(\alpha,\beta)}(\pi_D|\pi_C) \Rightarrow \text{BND}_{(1,0)}(\pi_D|\pi_R) = \text{BND}_{(1,0)}(\pi_D|\pi_C)$, and the reverse is not necessarily true. For a decision-theoretic rough set model, we can only get one-way inference, i.e.,

$$\overrightarrow{\text{POS}}_{(\alpha,\beta)}(\pi_D|\pi_R) = \overrightarrow{\text{POS}}_{(\alpha,\beta)}(\pi_D|\pi_C) \Rightarrow \text{POS}_{(\alpha,\beta)}(\pi_D|\pi_R) = \text{POS}_{(\alpha,\beta)}(\pi_D|\pi_C),$$

$$\overrightarrow{\text{BND}}_{(\alpha,\beta)}(\pi_D|\pi_R) = \overrightarrow{\text{BND}}_{(\alpha,\beta)}(\pi_D|\pi_C) \Rightarrow \text{BND}_{(\alpha,\beta)}(\pi_D|\pi_R) = \text{BND}_{(\alpha,\beta)}(\pi_D|\pi_C).$$

The reverse is not necessarily true. The regions based on vector representation do not have the monotonicity, as shown by the next example.

Example 2. (continued from Example 1) Let $\alpha = 0.40, \beta = 0.35$, $R = \{c_1\}$, we have:

$$\text{POS}_{(0.40,0.35)}(\pi_D|\pi_R) = \text{POS}_{(0.40,0.35)}(\pi_D|\pi_C) = U,$$

$$\overrightarrow{\text{POS}}_{(0.40,0.35)}(\pi_D|\pi_C) = (C_1, C_1 \cup C_3, C_2 \cup C_4),$$

$$\overrightarrow{\text{POS}}_{(0.40,0.35)}(\pi_D|\pi_R) = (C_1 \cup C_2, C_3 \cup C_4, U).$$

They show that the monotonicity does not hold.

4   A Reduct Construction Method

For an (α, β)-positive region attribute reduct, the positive rule set does not change if we merge equivalence classes with the same decisions. We can construct a discernibility matrix $M_{D_{POS}}([x]_C, [y]_C)$ for positive region attribute reduct as follows:

$$M_{D_{POS}}([x]_C, [y]_C) = \begin{cases} \{a \in C | I_a(x) \neq I_a(y)\}, \\ \quad \text{if } D_{POS(\alpha,\beta)}([x]_C) \neq D_{POS(\alpha,\beta)}([y]_C), \\ C, \quad \text{otherwise.} \end{cases} \quad (8)$$

By the discernibility function proposed by Skowron and Rauszer [11]:

$$f(M_{D_{POS}}) = \wedge \vee M_{D_{POS}}([x]_C, [y]_C). \quad (9)$$

We can get a decision reduct, which is a prime implicants of the reduced disjunctive form of the discernibility function.

If $\alpha > 0.5$, Eq. (8) degenerates into the discernibility matrix of a variable precision rough set (VPRS) model [8]. We can get an α-lower-approximation attribute reduct. Since VPRS is a special case of DTRS with $\alpha > 0.5$ and $\beta = 1 - \alpha$, we give a general conclusion for any pair of thresholds (α, β) with $0 \leq \beta < \alpha \leq 1$.

Theorem 2. *In a decision table DT, given a subset of attributes $A \subseteq C$ and a pair of thresholds $0 \leq \beta < \alpha \leq 1$, A is an (α, β)-positive region consistent set of C iff for any $x, y \in U$, $D_{POS(\alpha,\beta)}([x]_C) \neq D_{POS(\alpha,\beta)}([y]_C)$ implies $[x]_A \cap [y]_A = \emptyset$.*

Proof. ($\Rightarrow$) For any $x, y \in U$, if $[x]_A \cap [y]_A \neq \emptyset$, $[x]_A = [y]_A$. Thus, we have $D_{POS(\alpha,\beta)}([x]_A) = D_{POS(\alpha,\beta)}([y]_A)$. Since A is an (α, β)-positive region consistent set of C, $POS_{(\alpha,\beta)}(D_j|\pi_A) = POS_{(\alpha,\beta)}(D_j|\pi_C)$ for any $D_j \in \pi_D$. It follows that $x \in POS_{(\alpha,\beta)}(D_j|\pi_A) \Leftrightarrow x \in POS_{(\alpha,\beta)}(D_j|\pi_C)$, i.e., $D_{POS(\alpha,\beta)}([x]_A) = D_{POS(\alpha,\beta)}([x]_C)$. Similarly, we have $D_{POS(\alpha,\beta)}([y]_A) = D_{POS(\alpha,\beta)}([y]_C)$. Therefore, we can conclude that $D_{POS(\alpha,\beta)}([x]_C) = D_{POS(\alpha,\beta)}([y]_C)$, which conflicts with the assumption.

($\Leftarrow$) Since $A \subseteq C$, it is easy to verify that $T([x]_A) = \{[y]_C \mid [y]_C \in [x]_A/C\}$ is a partition of $[x]_A$.

(i) For any $D_j \in \pi_D$, if $x \in POS_{(\alpha,\beta)}(D_j|\pi_A)$, then $[x]_A \subseteq POS_{(\alpha,\beta)}(D_j|\pi_A)$. Since $[x]_A = [y]_A$ for any $y \in [x]_A$, according to the arbitrariness of y and the assumption that $D_{POS(\alpha,\beta)}([x]_C) = D_{POS(\alpha,\beta)}([y]_C)$, we have $D_j \in D_{POS(\alpha,\beta)}([x]_C)$, i.e., $x \in POS_{(\alpha,\beta)}(D_j|\pi_C)$.

(ii) If $x \in POS_{(\alpha,\beta)}(D_j|\pi_C)$, then $D_j \in D_{POS(\alpha,\beta)}([x]_C)$. Since $[x]_A = [y]_A$ for any $[y]_C \in T([x]_A)$, by the assumption, we have $D_{POS(\alpha,\beta)}([x]_C) = D_{POS(\alpha,\beta)}([y]_C)$. It follows that $D_j \in D_{POS(\alpha,\beta)}([y]_C)$ i.e., $\Pr(D_j|[y]_C) \geq \alpha$. Hence,

$$p(D_j|[x]_A) = \sum_{[y]_C \in T([x]_A)} \frac{|[y]_C|}{|[x]_A|} \frac{|[y]_C \cap D_j|}{|[y]_C|} > \alpha, \quad (10)$$

i.e., $x \in POS_{(\alpha,\beta)}(D_j|\pi_A)$.

Based on (i) and (ii), we have $\text{POS}_{(\alpha,\beta)}(D_j|\pi_A) = \text{POS}_{(\alpha,\beta)}(D_j|\pi_C)$ for any $D_j \in \pi_D$, which follows that A is an (α,β)-positive region consistent set of C.

According to Theorem 2, when deleting some attributes, if only equivalence classes with same decision set are merged, it can guarantee that the deleted attributes are redundant and the remaining set of attributes is an (α,β)-positive region consistent set of C.

Example 3. Consider a decision Table 3, where $U = \{x_1, \cdots, x_{10}\}$, $C = \{c_1, c_2\}$, and $D = \{d\}$. For the set of condition attributes, we have $U/C = \{C_1, C_2, C_3\}$, where $C_1 = \{x_1, x_2, x_3, x_4\}$, $C_2 = \{x_5, x_6, x_7\}$, $C_3 = \{x_8, x_9, x_{10}\}$. For the set of decision attributes, we have $U/D = \{D_1, D_2\}$, where $D_1 = \{x_1, x_2, x_3, x_7, x_8, x_9\}$ and $D_2 = \{x_4, x_5, x_6, x_{10}\}$. Let $R = \{c_1\}$. We have $U/R = \{R_1, R_2\}$, where $R_1 = C_1 \cup C_3$ and $R_2 = C_2$. If $\alpha = 0.6$, then $\overrightarrow{\text{POS}}_{(\alpha,\beta)}(\pi_D|\pi_C) = (C_1 \cup C_3, C_2)$ and $\overrightarrow{\text{POS}}_{(\alpha,\beta)}(\pi_D|\pi_R) = (R_1, R_2) = \overrightarrow{\text{POS}}_{(\alpha,\beta)}(\pi_D|\pi_C)$. It is easy to verity that $\{c_1\}$ is an (α,β)-positive region attribute reduct of C, it guarantees that the positive rules unchanged although their confidence is different. Before removing attributes, the positive rules are $C_1 \rightarrow_P D_1$, $C_3 \rightarrow_P D_1$, with confidence 0.750 and 0.667, respectively. By deleting attribute set $\{c_2\}$, equivalence classes C_1 and C_3 merge into R_1. We get a new positive rule $R_1 \rightarrow_P D_1$, with a confidence of 0.714.

On the other hand, we have $D_{\text{POS}(\alpha,\beta)}(C_1) = \{D_1\}$, $D_{\text{POS}(\alpha,\beta)}(C_2) = \{D_2\}$, $D_{\text{POS}(\alpha,\beta)}(C_3) = \{D_1\}$. The corresponding discernibility matrix $M_{\text{POS}(0.6,\beta)}$ is showed in Table 4. The attribute reduct is $\{c_1\}$, which is consistent with Definition 4.

Similarly, we can present the discernibility matrix $M_{D_{\neg\text{NEG}}}([x]_C, [y]_C)$ for non-negative region attribute reducts as follows:

$$M_{D_{\neg\text{NEG}}}([x]_C, [y]_C) = \begin{cases} \{a \in C | I_a(x) \neq I_a(y)\}, \\ \qquad \text{if } D_{\neg\text{NEG}(\alpha,\beta)}([x]_C) \neq D_{\neg\text{NEG}(\alpha,\beta)}([y]_C), \quad (11) \\ C, \quad \text{otherwise.} \end{cases}$$

Table 3. A decision table

U	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
c_1	1	1	1	1	0	0	0	1	1	1
c_2	1	1	1	1	0	0	0	0	0	0
d	1	1	1	2	2	2	1	1	1	2

Table 4. The discernibility matrix $M_{\text{POS}(0.6,\beta)}$ of Table 3

	C_1	C_2	C_3
C_1	C		
C_2	$\{c_1\}$	C	
C_3	C	$\{c_1, c_2\}$	C

Theorem 3. *Given a decision table DT, $A \subseteq C$ is an (α, β)-non-negative region consistent set of C iff for any $x, y \in U$, if $D_{\neg \text{NEG}(\alpha,\beta)}([x]_C) \neq D_{\neg \text{NEG}(\alpha,\beta)}([y]_C)$, then $[x]_A \cap [y]_A = \emptyset$.*

Proof. It is similar to the proof of theorem 2.

The discernibility function of the matrix is given by:

$$f(M_{D_{\neg \text{NEG}}}) = \wedge \vee M_{D_{\neg \text{NEG}}}([x]_C, [y]_C). \tag{12}$$

A prime implicant of the reduced disjunctive form of the discernibility function is an (α, β)-non-negative region attribute reduct.

5   Conclusion

In this paper, we analyze the decision region changes when deleting attributes. It is found that positive rules are unchanged when equivalence classes in positive region with the same decision set are merged. It is also true for the equivalence classes in non-negative region with the same decision set. The notion of vector based attribute reduct is proposed. One can use the standard method to construct a reduct based on a discernibility matrix.

References

1. Grzymala-Busse, J.W., Clarka, P.G., Kuehnhausena, M.: Generalized probabilistic approximations of incomplete data. Int. J. Approximate Reasoning **50**, 180–196 (2014)
2. Jia, X.Y., Liao, W.H., Tang, Z.M., Shang, L.: Minimum cost attribute reduction in decision-theoretic rough set models. Inf. Sci. **219**, 151–167 (2013)
3. Jia, X.Y., Shang, L., Zhou, B., Yao, Y.Y.: Generalized attribute reduct in rough set theory. Knowledge-Based Systems, 2015. http://dx.doi.org/10.1016/j.knosys.2015.05.017
4. Jia, X.Y., Tang, Z.M., Liao, W.H., Shang, L.: On an optimization representation of decision-theoretic rough set model. Int. J. Approximate Reasoning **55**, 156–166 (2014)
5. Li, H.X., Zhou, X.Z., Zhao, J.B., Liu, D.: Non-monotonic attribute reduction in decision-theoretic rough sets. Fundamenta Informaticae **126**, 415–432 (2013)
6. Liao, S.J., Zhu, Q.X., Min, F.: Cost-sensitive attribute reduction in decision-theoretic rough set models. Math. Prob. Eng. **2014**, 1–9 (2014)
7. Ma, X.A., Wang, G.Y., Yu, H., Li, T.R.: Decision region distribution preservation reduction in decision-theoretic rough set model. Inf. Sci. **278**, 614–640 (2014)
8. Mi, J.S., Wu, W.Z., Zhang, W.X.: Approaches to knowledge reduction based on variable precision rough set model. Inf. sci. **159**, 255–272 (2004)
9. Pawlak, Z.: Rough sets. Int. J. Comput. Inf. Sci. **11**, 341–356 (1982)

10. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht (1991)
11. Skowron, A., Rauszer, C.: The discernibility matrices and functions in information systems. In: Slowinski, R. (ed.) Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory, pp. 331–362. Kluwer, Dordrecht (1992)
12. Skowron, A.: Boolean reasoning for decision rules generation. In: Komorowski, J., Raś, Z.W. (eds.) Methodologies for Intelligent Systems. LNCS, vol. 689, pp. 295–305. Springer, Heidelberg (1993)
13. Yao, Y.Y.: Three-way decisions with probabilistic rough sets. Inf. Sci. **180**, 341–353 (2010)
14. Yao, Y.Y.: The two sides of the theory of rough sets. Knowl.-based Syst. **80**, 67–77 (2015)
15. Yao, Y.Y., Wong, S.K.M.: A decision theoretic framework for approximating concepts. Int. J. Man-Mach. Stud. **37**, 793–809 (1992)
16. Yao, Y.Y., Wong, S.K.M., Lingras, P.: A decision-theorectic rough set model. In: Ras, Z.W., Zemankova, M., Emrich, M.L. (eds.) Methodologies for Intelligent systems, vol. 5, pp. 17–24. Norh-Holland, New York (1990)
17. Yao, Y.Y., Zhao, Y.: Attribute reduction in decision-theoretic rough set models. Inf. Sci. **178**, 3356–3373 (2008)
18. Zhao, Y., Wong, S.K.M., Yao, Y.: A Note on Attribute Reduction in the Decision-Theoretic Rough Set Model. In: Peters, J.F., Skowron, A., Chan, C.-C., Grzymala-Busse, J.W., Ziarko, W.P. (eds.) Transactions on Rough Sets XIII. LNCS, vol. 6499, pp. 260–275. Springer, Heidelberg (2011)