Chapter 7
Structure-Based Attribute Reduction: A Rough Set Approach

Yoshifumi Kusunoki and Masahiro Inuiguchi

Abstract We provide an introduction to a rough set approach to attribute reduction. Analyzed data sets consist of objects which are described by attributes and partitioned into decision classes. Rough set theory deals with uncertainty decision classes with respect to attributes by approximating them to precise sets. The aim of attribute reduction is to remove redundant attributes as well as find important ones for classification. Several types of attribute reduction have been proposed especially according to preserving structures of approximated decision classes. We introduce definitions and theoretical results about structures-based attribute reduction.

Keywords Rough set model · Reduct · Boolean function · Structure-based reduct

7.1 Introduction

We provide an introduction to attribute reduction or feature selection based on rough set theory [35, 36, 39]. Rough set theory approaches uncertainty or inconsistency of membership for sets due to incomplete or granular information. In a rough set approach for data analysis, data sets are usually given by decision tables which consist of objects (items) described by attributes. Moreover, each object in decision tables is classified into decision classes. Because of incompleteness of given attributes, some objects are indiscernible to each other by the attributes, and that causes uncertainty of decision classes. Such an uncertain decision class is approximated by two precise sets, called lower and upper approximations. The difference of the upper and lower approximations is called a boundary.

Y. Kusunoki (✉)
Graduate School of Engineering, Osaka University, 2-1 Yamadaoka,
565-0871 Suita, Osaka, Japan
e-mail: kusunoki@eei.eng.osaka-u.ac.jp

M. Inuiguchi
Graduate School of Engineering Sciences, Osaka University, 1-3 Machikaneyama,
560-8531 Toyonaka, Osaka, Japan
e-mail: inuiguti@sys.es.osaka-u.ac.jp

One of the major topics for rough set based data analysis is (relative) attribute reduction [37, 39]. Attribute reduction is a problem to delete redundant condition (explanatory) attributes for the classification of the decision classes. Minimal sets of attributes preserving a part of information of the classification are called reducts. Reducts can be interpreted as important sets of attributes for the classification. Several types of reducts have been proposed according to a part of the information which should be preserved [23, 36, 43, 45]. Originally, Pawlak proposed reducts preserving the positive region [36, 43], which is the union of all lower approximations of decision classes, in other words, the set of all certainly classified objects. Ślęzak proposed ones preserving all boundaries of decision classes [45]. One of the authors also proposed two types of reducts, which preserve all lower approximations and all upper approximations of decision classes, and show that they are equivalent to reducts preserving the positive region and all boundaries, respectively [23].

Inspired by the above studies, we provide a framework to discuss attribute reduction in the rough set theory. We regard attribute reduction as removing condition attributes with preserving some part of the lower/upper approximations of the decision classes, because the approximations summarize the classification ability of the condition attributes. Hence, we define several types of reducts according to structures of the approximations [23, 24]. They are called "structure-based" reducts.

When several types of structure-based reducts are defined, we would be interested in whether one reduct is stronger/weaker than another reduct, in other words, one preserves more/less structure than the other. Therefore, we have investigated the strong-weak relation among different types of structure-based reducts. As a result of the investigation, we obtain a strong-weak hierarchy of structure-based reducts. The strong-weak hierarchy is useful when we search the best reduct for an application, because it provides a trade-off between the size (cost for precise classification) of a reduct and its classification ability. It is an advantage of the variations of structure-based reducts.

The rough set model is extended to apply to various kinds of data sets [12, 16, 22, 29, 38, 43, 47, 53, 54]. Two important extensions of the rough set model are the variable precision rough set model [53, 54] and the dominance-based rough set model [16]. The variable precision rough set model is a probabilistic extension. Given precision parameters, requirements for lower and upper approximations are relaxed to tolerate errors in decision tables. The dominance-based rough set model is applied to decision tables with ordinal attributes, where decision classes are ordered and monotonically depend on the ordinal attributes. It deals with inconsistency between the classification of the ordinal decision classes and the monotonic dependence. Instead of decision classes, upward unions and downward unions of decision classes are approximated. In the extended rough set models, we have studied structure-based reducts [20, 21, 25, 26, 31].

In the classical rough set model, it is well-known that reducts are associated with prime implicants of a Boolean function [37, 43]. We can efficiently enumerate reducts by converting it to enumerating prime implicants of the Boolean function. Like that conversion, the methodology solving a problem by solutions of a Boolean equation is called Boolean reasoning [37, 43]. In this chapter, we propose a unified

formulation of several Boolean functions corresponding to several types of reducts in the classical and extended rough set models.

In this chapter, we show our theoretical results of structure-based reducts in the classical and extended rough set models, including definitions of reducts and their strong-weak hierarchy. The results consist of our papers [20, 23, 25, 31]. Our main contributions are to propose structure-based reducts, investigate strong-weak relations of reducts, and connect reducts with prime implicants of Boolean functions in a unified formulation in the variable precision and dominance-based rough set models. For the structure-based reducts in the variable precision rough set model, we revise their definitions from our previous work [20]. Parts of the results were independently developed by other authors [33, 45, 49, 52].

This chapter is organized as follows. In Sect. 7.2, we study structure-based reducts in the classical rough set models. Firstly, we define a decision table and the rough set model of the decision table. Then, we introduce several types of reducts including structure-based reducts and others. We show that all types of reducts are reduced to two different types. Finally, we connect all reducts of each type with the prime implicants of a specific Boolean function. Sections 7.3 and 7.4 are devoted to structure-based reducts in the variable precision rough set model and those in the dominance-based rough set model, respectively. Those sections have almost the same organization as that of Sect. 7.2, namely, defining a rough set model and reducts, investigating strong-weak relations of reducts, and connecting reducts with prime implicants of Boolean functions. Concluding remarks are given in Sect. 7.5.

7.2 Structure-Based Attribute Reduction in Rough Set Models

7.2.1 Decision Tables

In rough set theory, analysed data sets form decision tables [36, 39]. A decision table is defined by $\mathbb{D} = (U, AT = C \cup \{d\}, \{V_a\}_{a \in AT})$.[1] U is a finite set of objects. AT is a finite set of attributes. V is a set of attribute values. Each attribute $a \in AT$ is a function $a : U \rightarrow V_a$, where $V_a \subseteq V$ is a set of values for a. For an object $u \in U$ and an attribute $a \in AT$, $a(u)$ is the value of u with respect to a. For $A = \{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\} \subseteq AT$, V_A is the Cartesian product of $\{V_{a_{i_l}}\}_{l=1,2,\ldots,k}$, namely, $V_A = \Pi_{a_{i_l} \in A} V_{a_{i_l}} = \{(v_{i_1}, v_{i_2}, \ldots, v_{i_k}) \mid v_{i_l} \in V_{a_{i_l}}, l = 1, 2, \ldots, k\}$. $A(u)$ is the tuple of the values of u with respect to A, namely, $A(u) = (a_{i_1}(u), a_{i_2}(u), \ldots, a_{i_k}(u))$. The attribute set AT is divided into a condition attribute set C and a decision attribute d to investigate the dependency of the decision attribute on condition attributes or the causal effect of condition attributes on the decision attribute. Throughout this chapter, we consider that the objects and the condition attributes are indexed by

---

[1] A decision table is often defined by the finite set of objects U and the finite set of attributes AT, i.e., (U, AT), however we use that definition to clarify the sets of values for the attributes.

$U = \{u_1, u_2, \ldots, u_n\}$ and $C = \{c_1, c_2, \ldots, c_m\}$, where $n = |U|$ and $m = |C|$. Moreover, we define the decision attribute values as $V_d = \{1, 2, \ldots, p\}$.

Remark 1 Decision tables are identical to data sets or data tables for the classification problem or the supervised learning in the data mining or machine learning literature, in which condition attributes are called attributes or independent variables, the decision attribute is a class attribute or dependent variable, and objects are tuples or samples. In that literature, each object is given by a tuple of attribute values with a class label (decision attribute value). However, we use the form of decision tables in this chapter by two reason. One is that we often deal with subsets of the attributes, so we prefer to let the symbols of the attributes be explicit. The other is to emphasise the view that a relation (e.g. the equivalence relation) on the object set is induced from a structure of the attribute value space (e.g. equivalence of values) through the attributes (functions).

Example 1 Consider a decision table $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ about car evaluations in Table 7.1, where $U = \{u_1, u_2, \ldots, u_7\}$, $C = \{\text{Pr, Ma, Sa}\}$ and $d = \text{Ev}$. The attribute value sets are given by $V_{\text{Pr}} = V_{\text{Ma}} = V_{\text{Sa}} = \{\text{low, med, high}\}$, $V_{\text{Ev}} = \{\text{unacc, acc, good}\}$. Condition attributes Pr, Ma, and Sa indicate price, maintenance cost, and safety of a car, respectively, by values high, med (medium), and low. Decision attribute Ev means evaluation of a car by some customer(s).

The value of u_1 with respect to Pr is $\text{Pr}(u_1) = \text{high}$, and that of u_2 with respect to Ev is $\text{Ev}(u_2) = \text{unacc}$. The value tuple of u_4 with respect to $C = \{\text{Pr, Ma, Sa}\}$ is $C(u_4) = (\text{med, high, low})$.

Given an attribute subset $A \subseteq AT$, we define an indiscernibility relation on U with respect to A, denoted by R_A, as follows:

$$R_A = \{(u, u') \in U^2 \mid a(u) = a(u'), \quad \text{for any } a \in A\}.$$

R_A is the set of the object pairs each of which is indiscernible by the given attributes A. Obviously, R_A is an equivalence relation, which is reflexive, symmetric, and transitive. From R_A, we define the equivalence class of an object $u \in U$, denoted by $R_A(u)$, as follows:

$$R_A(u) = \{u' \in U \mid (u', u) \in R_A\}.$$

Table 7.1 Decision table of car evaluations

Car	Pr	Ma	Sa	Ev
u_1	High	High	Low	Unacc
u_2	Med	Med	Med	Unacc
u_3	Med	Med	Med	Acc
u_4	Med	High	Low	Acc
u_5	Med	Med	High	Acc
u_6	Med	Med	High	Good
u_7	Low	Med	Med	Good

$R_A(u)$ is the set of objects which have the same values as u for all attributes in A. We denote the set of all equivalence classes with respect to R_A by $U/R_A = \{R_A(u) \mid u \in U\}$. Every equivalence class with respect to the decision attribute d is called a decision class. For each value of the decision attribute $i \in V_d$, we define the corresponding decision class $X_i = \{u \in U \mid d(u) = i\}$. Clearly, $\mathscr{X} = \{X_1, X_2, \ldots, X_p\}$ forms a partition of U.

Example 2 Remember $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ in Table 7.1. Let $A = \{\text{Pr}, \text{Ma}\}$ be an attribute subset. The discernibility relation R_A is described as the following matrix. Symbol $*$ indicates that the corresponding object pair u_i and u_j is in the discernibility relation, i.e., $(u_i, u_j) \in R_A$.

	u_1	u_2	u_3	u_4	u_5	u_6	u_7
u_1	*						
u_2		*	*		*	*	
u_3		*	*		*	*	
u_4				*			
u_5		*	*		*	*	
u_6		*	*		*	*	
u_7							*

From the matrix, we can easily see that the equivalence classes by R_A form a partition of U, namely, $U/R_A = \{\{u_1\}, \{u_4\}, \{u_7\}, \{u_2, u_3, u_5, u_6\}\}$.

The decision classes of the decision table $\mathbb{D}$ are obtained as $X_{\text{unacc}} = \{u_1, u_2\}$, $X_{\text{acc}} = \{u_3, u_4, u_5\}$, $X_{\text{good}} = \{u_6, u_7\}$.

7.2.2 Rough Set Models

Let A be a subset of the attribute set AT and X be a subset of the object set U. When X can be represented by a union of elements in U/R_A, we can say that the classification by X is consistent with the information of A. Such subsets of objects are called definable sets with respect to A. On the other hand, considering an object subset X which cannot be represented by any union of elements in U/R_A, the classification of X is inconsistent with A. The classical Rough Set Model (RSM) [35, 36, 39] deals with the inconsistency by two operators for object sets, called lower and upper approximations. For $A \subseteq AT$ and $X \subseteq U$, the lower approximation $\text{LA}_A(X)$ and the upper approximation $\text{UA}_A(X)$ of X with respect to A is defined by:

$$\text{LA}_A(X) = \{u \in U \mid R_A(u) \subseteq X\},$$
$$\text{UA}_A(X) = \{u \in U \mid R_A(u) \cap X \neq \emptyset\}.$$

The difference between $\text{UA}_A(X)$ and $\text{LA}_A(X)$ is called the boundary of X with respect to A, which is defined by:

$$\text{BN}_A(X) = \text{UA}_A(X) \setminus \text{LA}_A(X).$$

$\mathrm{LA}_A(X)$ is interpreted as the set of objects which are certainly classified to X in view of A. While, $\mathrm{UA}_A(X)$ is the set of objects which are possibly classified to X in view of A. $\mathrm{BN}_A(X)$ is a set of objects whose membership to X is doubtful.

The approximations are a definable set with respect to A, where a definable set with respect to A is a set defined by the union of elements in U/R_A:

$$\mathrm{LA}_A(X) = \bigcup_{R_A(u) \subseteq X} R_A(u) = \bigcup_{u \in \mathrm{LA}_A(X)} R_A(u),$$

$$\mathrm{UA}_A(X) = \bigcup_{R_A(u) \cap X \neq \emptyset} R_A(u) = \bigcup_{u \in \mathrm{UA}_A(X)} R_A(u).$$

The boundary is necessarily definable because U/R_A is the partition of U.
In fact, $\mathrm{LA}_A(X)$ and $\mathrm{UA}_A(X)$ are "lower" and "upper" approximations of X_i:

$$\mathrm{LA}_A(X) \subseteq X \subseteq \mathrm{UA}_A(X). \tag{7.1}$$

By the above inclusion relations and the definition of the boundary, it holds that

$$\mathrm{LA}_A(X) = X \setminus \mathrm{BN}_A(X), \tag{7.2}$$
$$\mathrm{UA}_A(X) = X \cup \mathrm{BN}_A(X). \tag{7.3}$$

For $B \subset A \subseteq AT$, we have,

$$\mathrm{LA}_B(X) \subseteq \mathrm{LA}_A(X) \text{ and } \mathrm{UA}_B(X) \supseteq \mathrm{UA}_A(X). \tag{7.4}$$

When B is included in A, the approximations with respect to B are coarser that those with respect to A. It means that dropping some attributes, i.e., information, decline the accuracy of RSM.

So far, we have defined approximations of X from the lower and upper definable sets. We can approximate the partition X and $U \setminus X$ by three definable sets. They are called positive, boundary, and negative regions of X with respect to A, denoted by $\mathrm{POS}_A(X)$, $\mathrm{BND}_A(X)$, and $\mathrm{NEG}_A(X)$, respectively:

$$\mathrm{POS}_A(X) = \bigcup \{E \in U/R_A \mid E \subseteq X\},$$
$$\mathrm{BND}_A(X) = \bigcup \{E \in U/R_A \mid E \cap X \neq \emptyset \text{ and } E \cap U \setminus X \neq \emptyset\},$$
$$\mathrm{NEG}_A(X) = \bigcup \{E \in U/R_A \mid E \subseteq U \setminus X\}.$$

$\mathrm{POS}_A(X)$ is the union of elements in U/R_A which are completely included in X, while $\mathrm{NEG}_A(X)$ is the union of elements in U/R_A which are completely excluded from X. $\mathrm{BND}_A(X)$ is the union of the rest of elements in U/R_A. Clearly, $\mathrm{POS}_A(X)$, $\mathrm{BND}_A(X)$, and $\mathrm{NEG}_A(X)$ form a partition of U. We can easily see the following correspondence:

$$\text{POS}_A(X) = \text{LA}_A(X),$$
$$\text{BND}_A(X) = \text{BN}_A(X),$$
$$\text{NEG}_A(X) = U \setminus \text{UA}_A(X).$$

In the rest of this section, we consider RSM for decision tables, namely, we only deal with approximations of decision classes $\mathscr{X} = \{X_1, X_2, \ldots, X_p\}$ with respect to subsets of condition attributes $A \subseteq C$.

Example 3 Remember the decision classes $X_{\text{unacc}} = \{u_1, u_2\}$, $X_{\text{acc}} = \{u_3, u_4, u_5\}$ and $X_{\text{good}} = \{u_6, u_7\}$ of the decision table in Table 7.1. The lower and upper approximations with respect to C of X_{unacc}, X_{cc} and X_{good} are obtained as follows:

$$\text{LA}_C(X_{\text{unacc}}) = \{u_1\}, \quad \text{UA}_C(X_{\text{unacc}}) = \{u_1, u_2, u_3\},$$
$$\text{LA}_C(X_{\text{acc}}) = \{u_4\}, \quad \text{UA}_C(X_{\text{acc}}) = \{u_2, u_3, u_4, u_5, u_6\},$$
$$\text{LA}_C(X_{\text{good}}) = \{u_7\}, \quad \text{UA}_C(X_{\text{good}}) = \{u_5, u_6, u_7\}.$$

We can see that $\text{LA}_C(X_i) \subseteq X_i \subseteq \text{UA}_C(X_i)$ for each $i = $ unacc, acc, good. Moreover, we can also see that each approximation is the union of equivalence classes included in the approximation, e.g., $\text{UA}_C(X_{\text{acc}}) = \{u_2, u_3\} \cup \{u_4\} \cup \{u_5, u_6\}$.

We reduce condition attributes to $A = \{\text{Pr}\}$. The approximations become:

$$\text{LA}_A(X_{\text{unacc}}) = \{u_1\}, \quad \text{UA}_A(X_{\text{unacc}}) = \{u_1, u_2, u_3, u_4, u_5, u_6\},$$
$$\text{LA}_A(X_{\text{acc}}) = \emptyset, \quad \text{UA}_A(X_{\text{acc}}) = \{u_2, u_3, u_4, u_5, u_6\},$$
$$\text{LA}_A(X_{\text{good}}) = \{u_7\}, \quad \text{UA}_A(X_{\text{good}}) = \{u_2, u_3, u_4, u_5, u_6, u_7\}.$$

The approximations with respect to A are coarser than those with respect to C, namely, $\text{LA}_A(X_i) \subseteq \text{LA}_C(X_i)$ and $\text{UA}_A(X_i) \supseteq \text{UA}_C(X_i)$ for each $i = $ unacc, acc, good.

For every X_i, the lower approximation $\text{LA}_A(X_i)$ and the boundary $\text{BN}_A(X_i)$ can be represented using all upper approximations of decision classes $\text{UA}_A(X_1)$, $\text{UA}_A(X_2), \ldots, \text{UA}_A(X_p)$:

$$\text{LA}_A(X_i) = \text{UA}_A(X_i) \setminus \bigcup_{j \in V_d \setminus \{i\}} \text{UA}_A(X_j), \tag{7.5}$$

$$\text{BN}_A(X_i) = \text{UA}_A(X_i) \cap \bigcup_{j \in V_d \setminus \{i\}} \text{UA}_A(X_j). \tag{7.6}$$

All upper approximations form a cover of U:

$$U = \bigcup_{i \in V_d} \text{UA}_A(X_i). \tag{7.7}$$

A positive region with respect to $A \subseteq C$ is also defined for the decision attribute d or equivalently for the decision table $\mathbb{D}$. It is the union of all positive regions of decision classes, i.e., the set of objects which are certainly classified to exactly one of the decision classes:

$$\text{POS}_A(d) = \bigcup_{i \in V_d} \text{POS}_A(X_i).$$

A generalized decision function [1, 45] with respect to $A \subseteq C$, denoted by $\partial_A : U \to 2^{V_d}$, provides a useful representation of RSM. For $u \in U$, $\partial_A(u)$ is a set of decision attribute values or decision classes to which u is possibly classified:

$$\partial_A(u) = \{i \in V_d \mid X_i \cap R_A(u) \neq \emptyset\}.$$

The generalized decision function gives an object-wise view of RSM. The lower and upper approximations can be expressed by the generalized decision function:

$$\text{LA}_A(X_i) = \{u \in U \mid \partial_A(u) = \{i\}\},$$
$$\text{UA}_A(X_i) = \{u \in U \mid \partial_A(u) \ni i\}.$$

Because $\partial_A(u)$ is defined based on $R_A(u)$, we have

$$\partial_A(u) = \partial_A(u') \text{ if } (u, u') \in R_A,$$

and because each object u is included in at least one upper approximation, we have

$$\partial_A(u) \neq \emptyset.$$

The monotonic property of upper approximations is represented as:

$$B \subseteq A \Rightarrow \partial_B(u) \supseteq \partial_A(u) \quad \text{for all } u \in U.$$

Example 4 Remember $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ in Table 7.1. The generalized decision function ∂_C is obtained as follows.

$\partial_C(u_1) = \{\text{unacc}\}, \quad \partial_C(u_2) = \partial_C(u_3) = \{\text{unacc, acc}\},$
$\partial_C(u_4) = \{\text{acc}\}, \quad \partial_C(u_5) = \partial_C(u_6) = \{\text{acc, good}\},$
$\partial_C(u_7) = \{\text{good}\}.$

For $A \subseteq C$, a quality of classification (or quality of approximation) of the decision attribute d with respect to A is defined by:

$$\gamma_A(d) = \frac{|\text{POS}_A(d)|}{|U|}. \tag{7.8}$$

It measures to what degree objects are correctly classified by RSM.

7.2.3 Reducts in Rough Set Models

7.2.3.1 Preserving Positive Region, Quality, and Generalized Decisions

Attribute reduction is to find important subsets of condition attributes by dropping as many as possible other condition attributes while preserving some specific information of RSM for a decision table $\mathbb{D}$. A minimal subset of condition attributes preserving the information is called a relative (or decision) reduct. In this chapter, we call it "reduct" for short. Reducts are originally defined to preserve the positive region $\text{POS}_C(d)$ [36, 43].

Definition 1 ([36, 43]) A reduct is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\text{POS}_A(d) = \text{POS}_C(d). \tag{P}$$

Here, the minimality is defined in terms of the set inclusion, i.e., there is no proper subset $A' \subset A$ satisfying (P).

A condition attribute subset A satisfying (P) preserves the information of the certain classification in the decision table. Generally, there exist more than one reduct in a decision table. The intersection of all reducts is called the core. Every element in the core is an essential condition attribute to preserve the information of $\text{POS}_C(d)$. The core can be empty. On the other hand, the condition attributes which do not belong to any reducts can be dropped without deterioration of the information. We call the original reduct a P-reduct.

Remark 2 Condition (P) is monotonic with respect to the set inclusion of condition attributes, i.e., for $A' \subseteq A \subseteq C$ we have $\text{POS}_{A'}(d) = \text{POS}_C(d)$ implies $\text{POS}_A(d) = \text{POS}_C(d)$. Hence, the above minimality condition for A is equivalent to that there is no condition attribute $a \in A$ such that $A \setminus \{a\}$ satisfies (P).

Example 5 Remember $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ in Table 7.1. The set of all condition attributes C obviously satisfies condition (P), but it is not a P-reduct because a proper subset $A = \{\text{Pr}, \text{Ma}\}$ satisfies (P). On the other hand, A is a P-reduct because all of proper subsets of A do not preserve the positive region: $\text{POS}_{\{\text{Pr}\}}(d) = \{u_1, u_7\}$ and $\text{POS}_{\{\text{Ma}\}}(d) = \text{POS}_\emptyset(d) = \emptyset$.

We can define another kind of reducts preserving the quality of classification [23, 39, 40].

Definition 2 ([40]) A Q-reduct is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\gamma_A(d) = \gamma_C(d). \tag{Q}$$

Clearly, condition (P) implies (Q). In RSM, the inverse is also true, because the monotonic property of POS(d) holds, namely, for $A' \subseteq A$, $\mathrm{POS}_{A'}(d) \subseteq \mathrm{POS}_A(d)$. We also call Q-reducts measure-based reducts, because it preserves a predefined measure for information of RSM.

Bazan et al. [1] and Ślęzak [45] proposed reducts preserving the generalized decision function ∂.

Definition 3 ([1, 45]) A G-reduct is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\partial_A(u) = \partial_C(u) \text{ for all } u \in U. \tag{G}$$

As shown in the next section, condition (G) implies (P).

7.2.3.2 Structure-Based Reducts

Structure-based reducts, proposed by one of the authors [20, 23], are defined to preserve families of object sets (structures) which are composed of lower and upper approximations, or positive, boundary, and negative regions. Hence, reducts of condition (P) can be seen as structure-based.

Now, we introduce structure-based reducts proposed in [23]. First, we define a reduct preserving all lower approximations. The preservation of the lower approximations implies the sustenance of certain classification ability.

Definition 4 ([23]) An L-reduct is a minimal condition attribute subset $A \subseteq C$ preserving the following condition:

$$\mathrm{LA}_A(X_i) = \mathrm{LA}_C(X_i) \text{ for all } i \in V_d. \tag{L}$$

Clearly, condition (L) implies (P) as well as (Q). In RSM, the inverse is also true, because the lower approximations $\mathrm{LA}(X_i)$, $i = 1, 2, \ldots, p$ have the empty intersection with each other, and they are monotonically decreasing with respect to the set inclusion of condition attributes.

However, even if we preserve lower approximations $\mathrm{LA}_C(X_i)$, $i = 1, 2, \ldots, p$, we may lose the information of boundaries $\mathrm{BN}_A(X_i)$, $i = 1, 2, \ldots, p$ and the information of upper approximations $\mathrm{UA}_A(X_i)$, $i = 1, 2, \ldots, p$.

Ślęzak [45] proposed a type of reducts preserving all boundaries.

Definition 5 ([45]) A B-reduct is a minimal condition attribute subset $A \subseteq C$ preserving the following condition:

$$\mathrm{BN}_A(X_i) = \mathrm{BN}_C(X_i) \quad \text{for all } i \in V_d. \tag{B}$$

The preservation of boundaries implies the protection against uncertainty expansion. Ślęzak [45] also showed that condition (B) is equivalent to (G). Hence, we have that a B-reduct is a G-reduct and vice versa.

On the other hand, we proposed a reduct preserving all upper approximations [23].

Definition 6 ([23]) A U-reduct is a minimal condition attribute subset $A \subseteq C$ preserving the following condition:

$$\mathrm{UA}_A(X_i) = \mathrm{UA}_C(X_i) \quad \text{for all } i \in V_d. \tag{U}$$

By definition, the classification ability of the upper approximations is equal to that of the generalized decision function. Hence, condition (U) is equivalent to (G), and we have that a U-reduct is a G-reduct and vice versa.

From Eqs. (7.2) and (7.5), we know that lower approximations are obtained from upper approximations as well as from boundaries. This fact implies that each of the preservation of all upper approximations or the preservation of all boundaries entails the preservation of all lower approximations.

To sum up the above discussion, we have the next theorem.

Theorem 1 ([23, 45]) *Let A be a subset of C. We have the following statements.*

(a) *A is a Q-reduct if and only if A is an L-reduct.*
(b) *A is a P-reduct if and only if A is an L-reduct.*
(c) *A is a G-reduct if and only if A is a U-reduct.*
(d) *A is a B-reduct if and only if A is a U-reduct.*
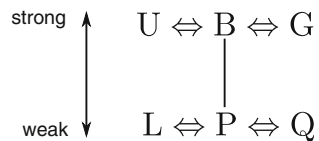(e) *A is a U-reduct as well as B-reduct, then A satisfies condition* (L).

All statements in the theorem can be easily proved by the equations which appeared in Sect. 7.2.2. For example, to prove Theorem 1(d), we show that preserving all boundaries implies preserving all upper approximations by Eq. (7.3), and show the converse by Eq. (7.6).

From Theorem 1(e), if A is a U-reduct then there exists an L-reduct $B \subseteq A$. Note that the converse is not always true, i.e., for an L-reduct B, there is no guarantee that there exists a U-reduct $A \supseteq B$.

The relations of 6 types of reducts are depicted in Fig. 7.1. Reducts located in the upper part of the figure preserve regions much more. Therefore, such reducts are larger in the sense of the set inclusion than the other reducts located in the lower part. A line segment connecting two types of reducts implies that, for each reduct of the upper type say A satisfies the preserving condition of the reduct of the lower one. From the figure, we know that there are 2 different types of reducts: U-reducts and L-reducts, and U-reducts are stronger than L-reducts.

Remark 3 As shown in Theorem 1, the six types of reducts are reduced to two types. However, it is important to define all possible types of reducts and organize them because of two reasons. One is that when we should mention different definitions of

Fig. 7.1 Strong-weak hierarchy of 6 types of reducts in RSM

strong $\uparrow$ $\mathrm{U} \Leftrightarrow \mathrm{B} \Leftrightarrow \mathrm{G}$

weak $\downarrow$ $\mathrm{L} \Leftrightarrow \mathrm{P} \Leftrightarrow \mathrm{Q}$

reducts (e.g. different authors would give different definitions), we can easily quote the equivalent of them from here. The other is that equivalent reducts (e.g. U-reducts and B-reducts) in RSM could become different in an extended RSM (e.g. variable precision RSM).

Remark 4 From the discussion above, we know that a U-reduct preserves more information than an L-reduct. However, when $p = 2$, we have the following relation:

$$\mathrm{UA}_A(X_1) = U \setminus \mathrm{LA}_A(X_2), \quad \mathrm{UA}_A(X_2) = U \setminus \mathrm{LA}_A(X_1).$$

Namely, we obtain upper approximations from lower approximations. Hence, in that case, an L-reduct is a U-reduct.

Remark 5 From Theorem 1, we see that preserving the measure γ is equivalent to preserving the lower approximations. Contrary, we can define a measure preserving which is equivalent to preserving the upper approximations. For example [23], we define,

$$\sigma_A(d) = \frac{\sum_{i \in V_d} |U \setminus \mathrm{UA}_A(X_i)|}{(p-1)|U|},$$

then $\sigma_A(d) = \sigma_C(d)$ is same as condition (U).

7.2.4 Boolean Functions Representing Reducts

Boolean reasoning [37] is a methodology where solutions of a given problem is associated with those of Boolean equations. In this section, we develop positive (monotone) Boolean functions whose solutions are given by condition attribute subsets satisfying the preserving conditions (L) or (U). Moreover, prime implicants of the Boolean functions exactly correspond to L-reducts or U-reducts. The Boolean functions are useful for enumerating reducts.

The results of this section are well-known and appeared in many papers e.g. [1, 43, 45, 50], but in slightly different expressions from ours. A unified formulation of Boolean functions of different types of reducts is provided using the generalized decision function.

Here, we briefly introduce Boolean functions and Boolean formulas [9, 14]. Let q be a natural number. A Boolean function is a mapping $f : \{0, 1\}^q \to \{0, 1\}$, where $w \in \{0, 1\}^q$ is called a Boolean vector whose ith component is w_i. Let $x_1, x_2, \ldots, x_q$ be Boolean variables. A Boolean formula in the Boolean variables $x_1, x_2, \ldots, x_q$ is a composition of 0, 1, the variables and operators of conjunction $\wedge$, disjunction $\vee$, complementation $\bar{\cdot}$, such as $x_1 \wedge (x_2 \vee \overline{x}_3)$, $(x_1 \wedge \overline{x}_2) \vee x_3$, and so on (for complete definition, see e.g. [9]). The Boolean formula is a Boolean function of the variables $x_1, x_2, \ldots x_q$. Conversely, any Boolean function can be expressed by a Boolean formula. For two Boolean functions f and g, $g \leq f$ means that f and g satisfy

$g(w) \leq f(w)$ for all $w \in \{0, 1\}^q$, and $g < f$ means that $g \leq f$ and $g \neq f$. A Boolean function f is positive or monotone, if $w \leq w'$ implies $f(w) \leq f(w')$ for all $w, w' \in \{0, 1\}^q$.

Boolean variables $x_1, x_2, \ldots$ and the complements $\overline{x}_1, \overline{x}_2, \ldots$ are called literals. A clause (resp., term) is a disjunction (resp., conjunction) of at most one of x_i and $\overline{x}_i$ for each variable. The empty disjunction (resp., conjunction) is denoted by $\perp$ (resp., $\top$). A clause c (resp., term t) is an implicate (resp., implicant) of a function f, if $f \leq c$ (resp. $t \leq f$). Moreover, it is prime if there is no implicate $c' < c$ (resp., no implicant $t' > t$) of f. A conjunction normal form (CNF) (resp., disjunction normal form (DNF)) of a function f is a Boolean formula of f which is expressed by a conjunction of implicates (resp. disjunction of implicants) of the function, and it is prime if all its members are prime. The complete CNF (resp. DNF) of f is the conjunction of all prime implicates (resp. disjunction of all prime implicants) of f. When f is positive, there is the unique CNF (resp. DNF) of f which is the complete CNF (resp. DNF) of f.

First, we associate conditions (L) (or (P)) and (U) (or (B)) with the conditions of the generalized decision function. As mentioned in the previous section, condition (U) is equivalent to (G).

Lemma 1 ([23, 50]) *Let A be a subset of C. We have the following statements.*

- *Condition* (L) *is equivalent to:*

$$\partial_A(u) = \partial_C(u) \quad \text{for all } u \in U \quad \text{such that } |\partial_C(u)| = 1. \qquad \text{(LG)}$$

- *Condition* (U) *is equivalent to* (G), *i.e.,*

$$\partial_A(u) = \partial_C(u) \quad \text{for all } u \in U. \qquad \text{(G)}$$

The next lemma is the heart of the Boolean reasoning, which connects two notions: "preserving" and "discerning".

Lemma 2 *For $u \in U$, the following assertions are equivalent.*

- $\partial_A(u) = \partial_C(u)$.
- $\forall u' \in U, \ (\partial_C(u') \neq \partial_C(u) \Rightarrow \exists a \in A, \ (u', u) \notin R_{\{a\}})$

Hence, to preserve the generalized decision of an object u, we should discern u from other objects u' having different generalized decisions from that of u.

Using Lemmas 1 and 2, we define two Boolean formulas, called discernibility functions. First, we define a discernibility matrix by $M = (m_{ij})_{i,j=1,2,\ldots,n}$, where ij-entry m_{ij} is a set of condition attributes which discern objects u_i and u_j,

$$m_{ij} = \{c \in C \mid c(u_i) \neq c(u_j)\}.$$

Then, we define discernibility functions.

Definition 7 Discernibility functions F^{U} and F^{L} are defined as follows:

$$F^{\mathrm{U}}(\tilde{c}_1, \ldots, \tilde{c}_m) = \bigwedge_{i,j \mid \partial_C(u_j) \neq \partial_C(u_i)} \bigvee_{c \in m_{ij}} \tilde{c},$$

$$F^{\mathrm{L}}(\tilde{c}_1, \ldots, \tilde{c}_m) = \bigwedge_{i \mid \; |\partial_C(u_i)|=1} \bigwedge_{j \mid \partial_C(u_j) \neq \partial_C(u_i)} \bigvee_{c \in m_{ij}} \tilde{c},$$

where $\tilde{c}_i$ is a Boolean variable corresponding to ith condition attribute c_i.

For $A \subseteq C$, we consider a Boolean vector $\tilde{c}^A = (\tilde{c}_1^A, \ldots, \tilde{c}_m^A)$, where,

$$\tilde{c}_i^A = \begin{cases} 1 & \text{if } c_i \in A, \\ 0 & \text{if } c_i \notin A. \end{cases}$$

Let $F^{\mathrm{U}}(\tilde{c}^A) = 1$. Then, for each pair u_i and u_j such that $\partial_C(u_i) \neq \partial_C(u_j)$, the intersection of A and m_{ij} should not be empty by the definition of F^{U}. By Lemma 2, in that case, $\partial_A(u) = \partial_C(u)$ for each u holds. We have the similar consequence when $F^{\mathrm{L}}(\tilde{c}^A) = 1$. Therefore, the following theorem holds. Let $\phi_A = \bigwedge \{\tilde{c} \mid c \in A\}$.

Theorem 2 ([43, 45, 50]) *Let A be a subset of C. We have the following equivalences:*

- *A satisfies* (G), *i.e.,* (U) *if and only if $F^{\mathrm{U}}(\tilde{c}^A) = 1$. Moreover, A is a U-reduct in RSM if and only if ϕ_A is a prime implicant of F^{U},*
- *A satisfies* (LG), *i.e.,* (L) *if and only if $F^{\mathrm{L}}(\tilde{c}^A) = 1$. Moreover, A is an L-reduct in RSM if and only if ϕ_A is a prime implicant of F^{L}.*

Definition 7 shows CNFs of F^{U} and F^{L}. The prime CNFs of the functions can be easily obtained. Because the functions are positive, the prime implicants of the prime DNF of each function are all of the prime implicants of the function. Therefore, all reducts of each type appear in the prime DNF of the corresponding function. The problem which converts the prime CNF of a positive Boolean function to its prime DNF is called the dualization problem [14]. We show an example for enumerating reducts by solving the dualization problems of the discernibility functions.

Example 6 Remember the decision table $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ in Table 7.1. In Table 7.2, we show again the decision table $\mathbb{D}$ with the generalized decision function ∂_C.

The discernibility matrix is obtained as below. Sign $*$ attached to objects u_i means that the generalized decision of u_i is a singleton, or equivalently, u_i is in the positive region $\mathrm{POS}_C(d)$.

Table 7.2 The decision table in Table 7.1 with the generalized decision function

Car	Pr	Ma	Sa	Ev	∂_C
u_1	High	High	Low	Unacc	{unacc}
u_2	Med	Med	Med	Unacc	{unacc, acc}
u_3	Med	Med	Med	Acc	{unacc, acc}
u_4	Med	High	Low	Acc	{acc}
u_5	Med	Med	High	Acc	{acc, good}
u_6	Med	Med	High	Good	{acc, good}
u_7	Low	Med	Med	Good	{good}

	u_1	u_2	u_3	u_4	u_5	u_6	u_7
u_1^*	Ø	C	C	{Pr}	C	C	C
u_2	C	Ø	Ø	{Ma, Sa}	{Sa}	{Sa}	{Pr}
u_3	C	Ø	Ø	{Ma, Sa}	{Sa}	{Sa}	{Pr}
u_4^*	{Pr}	{Ma, Sa}	{Ma, Sa}	Ø	{Ma, Sa}	{Ma, Sa}	C
u_5	C	{Sa}	{Sa}	{Ma, Sa}	Ø	Ø	{Pr, Sa}
u_6	C	{Sa}	{Sa}	{Ma, Sa}	Ø	Ø	{Pr, Sa}
u_7^*	C	{Pr}	{Pr}	C	{Pr, Sa}	{Pr, Sa}	Ø

The discernibility functions F^{L} and F^{L} for the decision table are calculated as:

$$F^{\mathrm{L}}(\tilde{\mathrm{Pr}}, \tilde{\mathrm{Ma}}, \tilde{\mathrm{Sa}}) = \bigwedge_{i=1,4,7} \bigwedge_{j \neq i} \bigvee_{c \in m_{ij}} \tilde{c} = (\tilde{\mathrm{Pr}}) \wedge (\tilde{\mathrm{Ma}} \vee \tilde{\mathrm{Sa}}) = (\tilde{\mathrm{Pr}} \wedge \tilde{\mathrm{Ma}}) \vee (\tilde{\mathrm{Pr}} \wedge \tilde{\mathrm{Sa}}),$$

$$F^{\mathrm{U}}(\tilde{\mathrm{Pr}}, \tilde{\mathrm{Ma}}, \tilde{\mathrm{Sa}}) = \bigwedge_{i,j|i \neq j,(i,j) \neq (2,3),(3,2),(5,6),(6,5)} \bigvee_{c \in m_{ij}} \tilde{c} = (\tilde{\mathrm{Pr}}) \wedge (\tilde{\mathrm{Sa}}) = (\tilde{\mathrm{Pr}} \wedge \tilde{\mathrm{Sa}}).$$

Therefore, there are two L-reducts {Pr, Ma} and {Pr, Sa}, and one U-reduct {Pr, Sa}.

In this case, we would select the U-reduct {Pr, Sa}, because we obtain the same size of reducts even if we select the other L-reduct.

7.3 Structure-Based Attribute Reduction in Variable Precision Rough Set Models

7.3.1 Rough Membership Function

The reason why decision tables are inconsistent is not only lack of knowledge (condition attributes) related to the decision attribute but also noise in observation of attribute values. In the latter case, the classical RSM would not be very useful because it does not permit any errors in the classification of objects into the lower approximations.

To overcome such shortcoming of the classical RSM, the variable precision rough set model (VPRSM) was proposed [53, 54]. Let $\mathbb{D} = (U, AT = C \cup \{d\}, \{V_a\}_{a \in AT})$ be a decision table. In definitions of lower and upper approximations in VPRSM, the following rough membership function of an object u with respect to an object set $X \subseteq U$ and an attribute set $A \subseteq AT$ plays an important role:

$$\mu_X^A(u) = \frac{|R_A(u) \cap X|}{|R_A(u)|}.$$

The value $\mu_X^A(u)$ gives the degree to which the object u belongs to the set X under the attribute set A. It can be interpreted as the conditional probability of $u \in X$ under $u \in R_A(u)$.

Because the rough membership function of an object is defined based not on the object but its equivalence class, we define a rough membership function of an equivalence class $E \in U/R_A$ for X:

$$\mu_X(E) = \frac{|E \cap X|}{|E|}.$$

An important property of the function is that given two equivalence classes $E, E' \in U/R_A$ the rough membership of the union $E \cup E'$ falls between those of E and E', namely,

$$\min\{\mu_X(E), \mu_X(E')\} \leq \mu_X(E \cup E') \leq \max\{\mu_X(E), \mu_X(E')\}. \tag{7.9}$$

7.3.2 Variable Precision Rough Set Models

Given precision parameters $0 \leq \beta < \alpha \leq 1$, lower and upper approximations of X with respect to A in VPRSM are defined as:

$$\mathrm{LA}_A^\alpha(X) = \{u \in U \mid \mu_X^A(u) \geq \alpha\},$$
$$\mathrm{UA}_A^\beta(X) = \{u \in U \mid \mu_X^A(u) > \beta\}.$$

The boundary of X is defined by $\mathrm{BN}_A^{\alpha,\beta}(X) = \mathrm{UA}_A^\beta(X) \setminus \mathrm{LA}_A^\alpha(X)$. When $\alpha = 1$ and $\beta = 0$, the approximations of X are the same as those of the classical RSM. $\mathrm{LA}_A^\alpha(X)$ is the set of objects whose degrees of membership to X are not less than α. On the other hand, $\mathrm{UA}_A^\beta(X)$ is the set of objects whose degrees of membership to X are more than β. In this chapter, we restrict our discussion to the situation that $\alpha = 1 - \beta$ and $\beta \in [0, 0.5)$. Under that situation, we have the dual property $\mathrm{LA}_A^\alpha(X) = U \setminus \mathrm{UA}_A^\beta(U \setminus X)$, because $\mu_X^A(u) = 1 - \mu_{U \setminus X}^A(u)$. We call β an admissible error rate. We denote $\mathrm{LA}_A^{1-\beta}(X)$ and $\mathrm{BN}_A^{1-\beta,\beta}(X)$ by $\mathrm{LA}_A^\beta(X)$ and $\mathrm{BN}_A^\beta(X)$, respectively.

Differently from (7.1) of RSM, we do not always have $\mathrm{LA}_A^\beta(X) \subseteq X$ and $\mathrm{UA}_A^\beta(X) \supseteq X$. However, we have

$$\mathrm{LA}_A^\beta(X) \subseteq \mathrm{UA}_A^\beta(X), \tag{7.10}$$

because $1 - \beta > \beta$ when $\beta < 0.5$. Moreover, we also have

$$\mathrm{LA}_A^\beta(X) \cap \mathrm{LA}_A^\beta(X') = \emptyset, \tag{7.11}$$

for any disjoint subsets $X, X' \subseteq U$, $X \cap X' = \emptyset$, because $\beta < 0.5$. Because the inclusion relation of (7.10), each of the lower and upper approximations, and the boundary is represented by the other two sets:

$$\mathrm{UA}_A^\beta(X) = \mathrm{LA}_A^\beta(X) \cup \mathrm{BN}_A^\beta(X),$$
$$\mathrm{LA}_A^\beta(X) = \mathrm{UA}_A^\beta(X) \setminus \mathrm{BN}_A^\beta(X).$$

The monotonic property (7.4) does not hold either. It causes difficulties of defining and enumerating reducts in VPRSM.

We can define positive, boundary, and negative regions in the same manner of the classical RSM:

$$\mathrm{POS}_A^\beta(X) = \bigcup \{R_A(u) \mid \mu_X^A(u) \geq 1 - \beta\},$$
$$\mathrm{BND}_A^\beta(X) = \bigcup \{R_A(u) \mid \mu_X^A(u) \in [\beta, 1 - \beta)\},$$
$$\mathrm{NEG}_A^\beta(X) = \bigcup \{R_A(u) \mid \mu_X^A(u) > \beta\}.$$

Clearly, we have,

$$\mathrm{POS}_A^\beta(X) = \mathrm{LA}_A^\beta(X),$$
$$\mathrm{BND}_A^\beta(X) = \mathrm{BN}_A^\beta(X),$$
$$\mathrm{NEG}_A^\beta(X) = U \setminus \mathrm{UA}_A^\beta(X).$$

In the rest of this section, we consider VPRSM under a decision table $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$. For each decision attribute value $i \in V_d$, the decision class $X_i = \{u \in U \mid d(u) = i\}$. The set of all decision classes are denoted by $\mathscr{X} = \{X_1, X_2, \ldots, X_p\}$.

Example 7 Consider a decision table $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ given in Table 7.3. The decision table composed of 40 objects with a condition attribute set $C = \{c_1, c_2, c_3, c_4\}$ and a decision attribute d. Each condition attribute takes a value bad or good, i.e., $V_{c_i} = \{\text{bad, good}\}$ for $i = 1, 2, 3, 4$. The decision attribute takes one of three values: $V_d = \{\text{bad, medium, good}\}$. Then there are three decision classes X_b, X_m and X_g whose objects take decision attribute value bad, medium and good,

Table 7.3 An example of decision table

	c_1	c_2	c_3	c_4	$d : (b, m, g)$
P_1	Good	Good	Bad	Good	(0, 2, 9)
P_2	Good	Good	Good	Bad	(0, 19, 1)
P_3	Bad	Good	Bad	Good	(1, 1, 2)
P_4	Bad	Bad	Bad	Good	(0, 1, 1)
P_5	Good	Bad	Good	Good	(1, 1, 1)

respectively. In Table 7.3, objects are classified into 5 groups $P_1, P_2, \ldots, P_5$ by the condition attributes C. For example, group P_1 is composed of objects having a condition attribute tuple $(c_1, c_2, c_3, c_4) = (\text{good, good, bad, good}) \in V_C$. The number of objects in each class in each group is shown in column $d : (b, m, g)$ in Table 7.3. For example, (0,2,9) of group P_1 means that no object is in class X_b, 2 objects are in class X_m and 9 objects are in class X_g.

The rough membership of an object u in each group P_i to each class X_k with respect to C is the number of objects in P_i and X_k divided by the number of objects in P_i. For example, $\mu_{X_b}^C(P_1) = 0/(0 + 2 + 9) = 0$, $\mu_{X_m}^C(P_1) = 2/(0 + 2 + 9) = 0.1818\ldots$, $\mu_{X_g}^C(P_1) = 9/(0 + 2 + 9) = 0.8181\ldots$. Given a condition attribute subset $A = \{c_1, c_2\}$, the objects in P_1 and P_2 are indiscernible to each other. Hence, the rough membership of an object u in P_1 and P_2 with respect to A becomes $\mu_{X_b}^A(P_1) = \mu_{X_b}^A(P_2) = 0/(0 + 21 + 10) = 0$, $\mu_{X_m}^A(P_1) = \mu_{X_m}^A(P_2) = 21/(0 + 21 + 10) = 0.6774\ldots$, $\mu_{X_g}^A(P_1) = \mu_{X_g}^A(P_2) = 10/(0 + 21 + 10) = 0.3225\ldots$.

Let $\beta = 0.39$. The lower approximations and the upper approximations with respect to C and β are obtained as follows:

$$\begin{aligned}
\mathrm{LA}_C^\beta(X_b) &= \emptyset, & \mathrm{UA}_C^\beta(X_b) &= \emptyset, \\
\mathrm{LA}_C^\beta(X_m) &= \{P_2\}, & \mathrm{UA}_C^\beta(X_m) &= \{P_2, P_4\}, \\
\mathrm{LA}_C^\beta(X_g) &= \{P_1\}, & \mathrm{UA}_C^\beta(X_g) &= \{P_1, P_3, P_4\},
\end{aligned}$$

where we express approximations by means of groups, namely, all members of a group P are members of an approximation X if $P \in X$.

In VPRSM, the properties corresponding to (7.5) and (7.6) are not always satisfied. Consequently, L-reducts, U-reducts, and B-reducts become independent concepts in VPRSM, and there are no strong-weak relations among them.

Additionally, property (7.7) only partially holds:

$$\frac{1}{p} > \beta \Rightarrow U = \bigcup_{i \in V_d} \mathrm{UA}_A^\beta(X_i). \tag{7.12}$$

The union of upper approximations of all decision classes does not always equal to U but when $1/p > \beta$. From this fact we define an unpredictable region of d with respect to β and A, denoted by $\mathrm{UNP}_A^\beta(d)$, as follows:

$$\mathrm{UNP}_A^\beta(d) = \bigcap_{i \in V_d} \mathrm{NEG}_A^\beta(X_i),$$

equivalently,

$$\text{UNP}_A^{\beta}(d) = U - \bigcup_{i \in V_d} \text{UA}_A^{\beta}(X_i).$$

The unpredictable region is the set of all objects which cannot be classified to any decision class.

We can define the positive region of d with respect to β and A in the same manner of RSM,

$$\text{POS}_A^{\beta}(d) = \bigcup_{i \in V_d} \text{POS}_A^{\beta}(X_i).$$

The quality of classification of d can be also defined in the same manner,

$$\gamma_A^{\beta}(d) = \frac{|\text{POS}_A^{\beta}(d)|}{|U|}.$$

The generalized decision function in RSM can be extended in VPRSM. However, differently from RSM, we define two functions. They are called lower and upper generalized decision functions, denoted by λ and υ, respectively. For each $u \in U$,

$$\lambda_A^{\beta}(u) = \{i \in V_d \mid \mu_{X_i}^A(u) \geq 1 - \beta\},$$
$$\upsilon_A^{\beta}(u) = \{i \in V_d \mid \mu_{X_i}^A(u) > \beta\}.$$

The lower generalized decision of u is the set of the decision values to which the membership degree of u is more than or equal to $1 - \beta$. The upper generalized decision of u is the set of the decision values to which the membership degree of u is more than β. The upper generalized decision corresponds to the generalized decision in RSM. By the definitions, the lower and upper generalized decision functions are closely related to the lower and upper approximations,

$$i \in \lambda_A^{\beta}(u) \Leftrightarrow u \in \text{LA}_A^{\beta}(X_i),$$
$$i \in \upsilon_A^{\beta}(u) \Leftrightarrow u \in \text{UA}_A^{\beta}(X_i).$$

So, they have the inclusion relation:

$$\lambda_A^{\beta}(u) \subseteq \upsilon_A^{\beta}(u). \tag{7.13}$$

Any two objects in the same equivalence class take the same values of generalized decision functions.

For each $(u, u') \in R_A$, $\lambda_A^{\beta}(u) = \lambda_A^{\beta}(u')$ and $\upsilon_A^{\beta}(u) = \upsilon_A^{\beta}(u')$. $\tag{7.14}$

The lower generalized decision function is a singleton or the empty set,

$$|\lambda_A^\beta(u)| \leq 1. \tag{7.15}$$

Unlike RSM, we may have the case that the upper generalized decision of u is a singleton, i.e., $\upsilon_A^\beta(u) = \{i\}$ but u does not belong to the lower approximation $\mathrm{LA}_A^\beta(X_i)$. When the lower generalized decision is a singleton, the upper generalized decision is also a singleton, and they are the same,

$$|\lambda_A^\beta(u)| = 1 \Rightarrow \upsilon_A^\beta(u) = \lambda_A^\beta(u). \tag{7.16}$$

Property (7.12) can be expressed as:

$$\frac{1}{p} > \beta \Rightarrow \upsilon_A^\beta(u) \neq \emptyset. \tag{7.17}$$

Hence, $\upsilon_A^\beta(u)$ may be empty unless β is less than $\frac{1}{p}$.

We define a function $(\upsilon\backslash\lambda)_A^\beta(u)$ as:

$$(\upsilon\backslash\lambda)_A^\beta(u) = \upsilon_A^\beta(u) \setminus \lambda_A^\beta(u).$$

By properties (7.13), (7.15), and (7.16), we have

$$(\upsilon\backslash\lambda)_A^\beta(u) = \emptyset \Rightarrow \upsilon_A^\beta(u) = \emptyset \text{ or } \lambda_A^\beta(u) \neq \emptyset, \tag{7.18}$$

$$(\upsilon\backslash\lambda)_A^\beta(u) \neq \emptyset \Rightarrow (\upsilon\backslash\lambda)_A^\beta(u) = \upsilon_A^\beta(u). \tag{7.19}$$

By that property, the following equivalence holds:

$$i \in (\upsilon\backslash\lambda)_A^\beta(u) \Leftrightarrow u \in \mathrm{BN}_A^\beta(X_i). \tag{7.20}$$

Therefore, we call $(\upsilon\backslash\lambda)$ a boundary generalized decision function.

Example 8 Remember the decision table $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ in Table 7.3. Let $\beta = 0.39$. The lower and upper generalized decision function with respect to C and β are,

$$\lambda_C^\beta(P_1) = \{\mathrm{g}\}, \lambda_C^\beta(P_2) = \{\mathrm{m}\}, \lambda_C^\beta(P_3) = \emptyset, \quad \lambda_C^\beta(P_4) = \emptyset, \quad \lambda_C^\beta(P_5) = \emptyset,$$
$$\upsilon_C^\beta(P_1) = \{\mathrm{g}\}, \upsilon_C^\beta(P_2) = \{\mathrm{m}\}, \upsilon_C^\beta(P_3) = \{\mathrm{g}\}, \upsilon_C^\beta(P_4) = \{\mathrm{m,g}\}, \upsilon_C^\beta(P_5) = \emptyset,$$

where $\lambda_C^\beta(P_i)$ and $\upsilon_C^\beta(P_i)$ indicate the lower and upper generalized decisions of an object in the group P_i.

7.3.3 Structure-Based Reducts in Variable Precision Rough Set Models

Before we define structure-based reducts in VPRSM, we firstly introduce Q-reducts. They preserve the quality of classification with the parameter β.

Definition 8 ([4, 5]) Let $\beta \in [0, 0.5)$ be an admissible error rate. A Q-reduct with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following conditions:

$$\gamma_A^\beta(d) = \gamma_C^\beta(d), \tag{VPQ1}$$

$$B \text{ satisfies (VPQ1)} \quad \text{for all} \quad B \supseteq A. \tag{VPQ2}$$

Remark 6 In VPRSM, approximations are no longer monotonic with respect to the set inclusion of condition attributes. Hence, condition (VPQ1) is not monotonic with respect to condition attributes, namely, A satisfies (VPQ1) but $B \supset A$ does not. In that case, we modify the preserving condition of reducts by adding a condition like (VPQ2). We notice that Beynon [4, 5] originally proposed Q-reducts (the author called them β-reducts) using only (VPQ1).

We define 4 kinds of structure-based reducts in VPRSM [20], which are already discussed in the classical RSM.

Definition 9 ([20, 33]) Let $\beta \in [0, 0.5)$ be an admissible error rate.

- A P-reduct[2] with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following conditions:

$$\text{POS}_A^\beta(d) = \text{POS}_C^\beta(d), \tag{VPP1}$$

$$B \text{ satisfies (VPP1)} \quad \text{for all} \quad B \supseteq A. \tag{VPP2}$$

- An L-reduct with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\text{LA}_A^\beta(X_i) = \text{LA}_C^\beta(X_i) \quad \text{for all} \ i \in V_d. \tag{VPL}$$

- A B-reduct with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following conditions:

$$\text{BN}_A^\beta(X_i) = \text{BN}_C^\beta(X_i) \quad \text{for all } i \in V_d, \tag{VPB1}$$

$$B \text{ satisfies (VPB1)} \quad \text{for all} \quad B \supseteq A. \tag{VPB2}$$

---

[2] Strictly speaking, P-reducts do not appear in [20].

- A U-reduct with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{UA}_A^\beta(X_i) = \mathrm{UA}_C^\beta(X_i) \qquad \text{for all } i \in V_d. \qquad \text{(VPU)}$$

Mi et al. [33] independently proposed L-reducts and U-reducts under the names of lower and upper distribution reducts.

Additionally, we can define a reduct preserving the unpredictable region.

Definition 10 ([20]) Let $\beta \in [0, 0.5)$ be an admissible error rate. A UN-reduct with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following conditions:

$$\mathrm{UNP}_A^\beta(d) = \mathrm{UNP}_C^\beta(d), \qquad \text{(VPUN1)}$$

$$B \text{ satisfies (VPUN1) for all } B \supseteq A. \qquad \text{(VPUN2)}$$

Remark 7 We modify the definitions of B- and UN-reducts from our paper [20], because there are mistakes in Boolean functions for B- and UN-reducts. By adding the second condition, the preserving conditions of B- and UN-reducts become monotone with respect to the set-inclusion of condition attribute sets.

By definitions, (VPL) and (VPU) obviously imply (VPP1,2) and (VPUN1,2), respectively. Moreover, (VPP1,2) also implies (VPQ1,2). Hence, we have the following relations among different types of reducts.

Theorem 3 ([20]) *Let A be a subset of C. We have the following statements in VPRSM with a fixed parameter* $\beta \in [0, 0.5)$,

(a) *A is an L-reduct then A satisfies (VPP1,2),*
(b) *A is a U-reduct then A satisfies (VPUN1,2),*
(c) *A is a P-reduct then A satisfies (VPQ1,2).*

Contrary to the classical RSM, (VPB1,2) is not equivalent to (VPU). In RSM, preserving boundaries implies preventing ambiguity expansion, namely upper approximations. However, in VPRSM, the ambiguity expansion can be prevented not only by preserving boundaries but by preserving them with the unpredictable region. Furthermore, we can define other compositions of different types of reducts.

Simply combining 5 types of structure-based reducts, we obtain $2^5 - 1 = 31$ types of reducts (ignoring (a) and (b) of Theorem 3). To reduce the number, we first investigate relationships of preserving conditions of reducts.

Theorem 4 ([20]) *Let A be a subset of C. We have the following statements in VPRSM with a fixed parameter* $\beta \in [0, 0.5)$,

- *The conjunction of* (VPB1) *and* (VPP1) *is equivalent to that of* (VPB1) *and* (VPUN1),

- *The conjunction of* (VPL) *and* (VPB1) *is equivalent to that of* (VPL) *and* (VPU),
- *The conjunction of* (VPU) *and* (VPP1) *is equivalent to that of* (VPL) *and* (VPU).

We define 4 different types of reducts.

Definition 11 ([20]) Let $\beta \in [0, 0.5)$ be an admissible error rate.

- An LU-reduct with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{LA}_A^\beta(X_i) = \mathrm{LA}_C^\beta(X_i) \text{ and } \mathrm{UA}_A^\beta(X_i) = \mathrm{UA}_C^\beta(X_i) \quad \text{for all } i \in V_d. \quad \text{(VPLU)}$$

- An LUN-reduct with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following conditions:

$$\mathrm{LA}_A^\beta(X_i) = \mathrm{LA}_C^\beta(X_i) \quad \text{for all } i \in V_d, \text{ and } \mathrm{UNP}_A^\beta(d) = \mathrm{UNP}_C^\beta(d),$$
$$\text{(VPLUN1)}$$

$$B \text{ satisfies (VPLUN1)} \quad \text{for all } B \supseteq A. \quad \text{(VPLUN2)}$$

- A BUN-reduct with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following conditions:

$$\mathrm{BN}_A^\beta(X_i) = \mathrm{BN}_C^\beta(X_i) \quad \text{for all } i \in V_d, \text{ and } \mathrm{UNP}_A^\beta(d) = \mathrm{UNP}_C^\beta(d),$$
$$\text{(VPBUN1)}$$

$$B \text{ satisfies (VPBUN1)} \quad \text{for all } B \supseteq A. \quad \text{(VPBUN2)}$$

- A PUN-reduct with β in VPRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following conditions:

$$\mathrm{POS}_A^\beta(d) = \mathrm{POS}_C^\beta(d) \quad \text{and } \mathrm{UNP}_A^\beta(d) = \mathrm{UNP}_C^\beta(d), \quad \text{(VPPUN1)}$$
$$B \text{ satisfies (VPPUN1)} \quad \text{for all } B \supseteq A. \quad \text{(VPPUN2)}$$

In Fig. 7.2, we show the relationships among 9 types of reducts. Names of reducts are abbreviated to their first characters. Reducts located in the upper part of Fig. 7.2 preserve regions much more. Therefore, such reducts are larger in the sense of set inclusion than the other reducts located in the lower part. A line segment connecting two types of reducts implies that, for each reduct of the upper type say A satisfies the preserving condition of a reduct of the lower one. From Fig. 7.2, we know that LU-reducts preserve regions most. On the other hand, UN-reducts and P-reducts do not preserve many regions.

The next proposition says that composite reducts such as LUN- or BUN-reducts can be constructed from their base reducts such as L- and UN-reducts or B- and UN-reducts. The proposition is useful to enumerate the composite reducts.
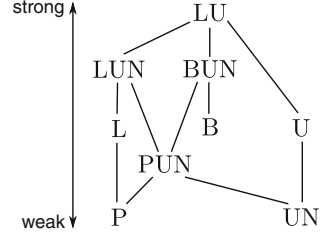
Proposition 1 *Consider two types of reducts, ♡-reducts and ♠-reducts, and the composition of them: ♡♠-reducts. Let $\mathscr{H}$ and $\mathscr{S}$ be the set of all ♡-reducts and the set of all ♠-reducts, respectively. Then the set of all ♡♠-reducts is the set of all minimal elements of $\{A \cup B \mid A \in \mathscr{H} \text{ and } B \in \mathscr{S}\}$.*

7.3.4 Boolean Functions Representing Reducts

As shown above, L- and U-reducts in the classical RSM are characterized by prime implicants of certain Boolean functions. In this section, we discuss Boolean functions of 9 types of reducts in VPRSM. To do this, we focus on Boolean functions of reducts pertaining to the lower approximations, the upper approximations, the boundaries, the positive region, and the unpredictable region, since the others can be obtained by taking conjunctions of those Boolean functions or using Proposition 1.

First, we represent the preserving conditions by the generalized decision functions.

Lemma 3 *Let $\beta \in [0, 0.5)$ be an admissible error rate, and A be a subset of C. We have the following statements:*

- *Condition (VPL) with β is equivalent to:*

$$\lambda_A^\beta(u) = \lambda_C^\beta(u) \quad \text{for all } u \in U. \tag{VPLG}$$

- *Condition (VPU) with β is equivalent to:*

$$\upsilon_A^\beta(u) = \upsilon_C^\beta(u) \quad \text{for all } u \in U. \tag{VPUG}$$

- *Condition (VPB1) with β is equivalent to:*

$$(\upsilon \backslash \lambda)_A^\beta(u) = (\upsilon \backslash \lambda)_C^\beta(u) \quad \text{for all } u \in U. \tag{VPBG1}$$

- *Condition (VPP1) with β is equivalent to:*

$$\lambda_A^\beta(u) = \emptyset \Leftrightarrow \lambda_C^\beta(u) = \emptyset \quad \text{for all } u \in U. \tag{VPPG1}$$

- *Condition (VPUN1) with β is equivalent to:*

$$\upsilon_A^\beta(u) = \emptyset \Leftrightarrow \upsilon_C^\beta(u) = \emptyset \quad \text{for all } u \in U. \tag{VPUNG1}$$

The next lemma is the counterpart of Lemma 2 of RSM. However, only the sufficient condition of the lemma holds in VPRSM.

Lemma 4 *Let $u \in U$ be an object, $\beta \in [0, 0.5)$ be an admissible error rate, and A be a subset of C.*

- *The following assertion is a sufficient condition of $\upsilon_A^\beta(u) = \upsilon_C^\beta(u)$:*

$$\forall u' \in U, \ (\upsilon_C^\beta(u') \neq \upsilon_C^\beta(u) \Rightarrow \exists a \in A, \ (u', u) \notin R_{\{a\}}).$$

- *The following assertion is a sufficient condition of $\lambda_A^\beta(u) = \lambda_C^\beta(u)$:*

$$\forall u' \in U, \ (\lambda_C^\beta(u') \neq \lambda_C^\beta(u) \Rightarrow \exists a \in A, \ (u', u) \notin R_{\{a\}}).$$

This lemma holds due to property (7.9). Then, we have the following corollary.

Corollary 1 *We have the following equivalences:*

$$\forall u \in U, \ \upsilon_A^\beta(u) = \upsilon_C^\beta(u)$$
$$\Leftrightarrow \forall u, u' \in U, \ (\upsilon_C^\beta(u') \neq \upsilon_C^\beta(u) \Rightarrow \exists a \in A, \ (u', u) \notin R_{\{a\}}),$$
$$\forall u \in U, \ \lambda_A^\beta(u) = \lambda_C^\beta(u)$$
$$\Leftrightarrow \forall u, u' \in U, \ (\lambda_C^\beta(u') \neq \lambda_C^\beta(u) \Rightarrow \exists a \in A, \ (u', u) \notin R_{\{a\}}).$$

It says that all L-reducts and all U-reducts can be enumerated by discernibility functions. The similar result is shown in [33]. However, we do not have the same result for $(\upsilon \backslash \lambda)$ and conditions (VPPG1) and (VPUNG1).

We introduce a discernibility matrix $M = (m_{i,j})_{ij=1,2,\dots,n}$, where ij-entry m_{ij} is defined by:

$$m_{ij} = \{c \in C \mid a(u_i) \neq a(u_j)\}.$$

It is the same as that of RSM. Then, we define discernibility functions corresponding to L-reducts and U-reducts, which are denoted by F_β^U and F_β^L, respectively.

Definition 12 Let $\beta \in [0, 0.5)$ be an admissible error rate. Discernibility functions F_β^U and F_β^L are defined as follows:

$$F_\beta^U(\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_m) = \bigwedge_{i,j \ | \ \upsilon_C^\beta(u_i) \neq \upsilon_C^\beta(u_j)} \bigvee_{c \in m_{ij}} \tilde{c},$$

$$F_\beta^{\mathrm{L}}(\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_m) = \bigwedge_{i,j \mid \lambda_C^\beta(u_i) \neq \lambda_C^\beta(u_j)} \bigvee_{c \in m_{ij}} \tilde{c},$$

where, $\tilde{c}_i$ is a Boolean variable pertaining to a condition attribute $c_i \in C$.

Function F_β^{U} is true if and only if at least one variable $\tilde{c}$ in m_{ij} of $\upsilon_C^\beta(u_i) \neq \upsilon_C^\beta(u_j)$ is true. While function F_β^{L} is true if and only if at least one variable $\tilde{c}$ in m_{ij} of $\lambda_C^\beta(u_i) \neq \lambda_C^\beta(u_j)$ is true.

Remember that we associate $A \subseteq C$ with a Boolean vector $\tilde{c}^A = (\tilde{c}_1^A, \tilde{c}_2^A, \ldots, \tilde{c}_m^A)$ as follows:

$$\tilde{c}_k^A = \begin{cases} 1 & c_k \in A, \\ 0 & \text{otherwise.} \end{cases}$$

Then, we can prove the next theorem from Corollary 1. Remember that ϕ_A is the term $\bigwedge\{\tilde{a} \mid a \in A\}$.

Theorem 5 ([20, 33]) *Let A be the subset of C, and $\beta \in [0, 0.5)$ be an admissible error rate. We have the following equivalences:*

- *A satisfies (VPUG) as well as (VPU) with β if and only if $F_\beta^{\mathrm{U}}(\tilde{c}^A) = 1$. Moreover, A is a U-reduct with β if and only if ϕ_A is a prime implicant of F_β^U,*
- *A satisfies (VPLG) as well as (VPL) with β if and only if $F_\beta^{\mathrm{L}}(\tilde{c}^A) = 1$. Moreover, A is an L-reduct with β if and only if ϕ_A is a prime implicant of F_β^{L}.*

For the preservation of the boundaries, the positive region, and the unpredictable region, we cannot use discernibility function approach. Because we cannot obtain a preserving subset $A \subseteq C$ by determining which pairs of objects should be discerned. For example, consider a decision table below.

	c_1	c_2	c_3	X_1	X_2	X_3
P_1	0	0	0	4	0	0
P_2	0	0	1	0	2	0
P_3	0	1	0	0	0	1
P_4	1	1	0	1	1	1

There are 3 condition attributes $C = \{c_1, c_2, c_3\}$ with the value set $V = \{0, 1\}$, and 3 decision classes X_1, X_2, X_3. P_1, P_2, P_3, P_4 are sets of objects, where members of each set have the same condition attribute values. The distribution of the decision classes on each set P_i is shown in the table, for instance the distribution on P_1 forms $|X_1 \cap P_1| = 4$, $|X_2 \cap P_1| = 0$, and $|X_3 \cap P_1| = 0$. Consider P-reducts with $\beta = 0.4$. The positive region of the table is $\mathrm{POS}_C^\beta(d) = P_1 \cup P_2 \cup P_3$. When we can make P_1 and P_2 be indiscernible and combine $P_1 \cup P_2$, the positive region is still preserved. Because the distribution on $P_1 \cup P_2$ is $(X_1, X_2, X_3) = (4, 2, 0)$, and $\mu_{X_1}(P_1 \cup P_2) =$

$2/3 \geq 0.6$. Similarly, we can make the pair of P_1 and P_3 and the pair of P_2 and P_3 be indiscernible. However, when we make all of P_1, P_2, and P_3 indiscernible, and select $\{c_1\}$ as a reduct, $P_1 \cup P_2 \cup P_3$ falls outside of $\mathrm{POS}_{\{c_1\}}^{\beta}(d)$, because the distribution is $(X_1, X_2, X_3) = (4, 2, 1)$, and $\mu_{X_1}(P_1 \cup P_2 \cup P_3) = 4/7 \leq 0.6$.

To overcome that difficulty, for each type of reducts, we consider two approximate discernibility functions $\hat{F}_\beta$ and $\check{F}_\beta$: $\hat{F}_\beta$ characterizes a sufficient condition of the preservation and $\check{F}_\beta$ characterizes a necessary condition.

First, we discuss discernibility functions characterizing sufficient conditions. By Theorems 3 and 4, we know that F_β^{L}, F_β^{U}, and $F_\beta^{\mathrm{LU}} = F_\beta^{\mathrm{L}} \wedge F_\beta^{\mathrm{U}}$ are discernibility functions of sufficient conditions for B-reducts, P-reducts, and UN-reducts with β.

Definition 13 Let $\beta \in [0, 0.5)$ be an admissible error rate. Discernibility functions $\hat{F}_\beta^{\mathrm{B}}$, $\hat{F}_\beta^{\mathrm{P}}$, and $\hat{F}_\beta^{\mathrm{UN}}$ are defined as follows:

$$\hat{F}_\beta^{\mathrm{B}} = F_\beta^{\mathrm{L}} \wedge F_\beta^{\mathrm{U}}, \quad \hat{F}_\beta^{\mathrm{P}} = F_\beta^{\mathrm{L}}, \quad \hat{F}_\beta^{\mathrm{UN}} = F_\beta^{\mathrm{U}}.$$

Clearly, we have the following proposition.

Proposition 2 ([20]) *Let A be a subset of C, and $\beta \in [0, 0.5)$ be an admissible error rate. We have the following implications:*

- *If $\hat{F}_\beta^{\mathrm{B}}(\tilde{c}^A) = 1$ then A satisfies (VPB1) and (VPB2) with β,*
- *If $\hat{F}_\beta^{\mathrm{P}}(\tilde{c}^A) = 1$ then A satisfies (VPP1) and (VPP2) with β,*
- *If $\hat{F}_\beta^{\mathrm{UN}}(\tilde{c}^A) = 1$ then A satisfies (VPUN1) and (VPUN2) with β.*

Next, let us discuss a discernibility function characterizing a necessary condition. Consider necessary discernibility functions for P-reducts. In the sufficient discernibility function $\hat{F}_\beta^{\mathrm{P}} = F_\beta^{\mathrm{L}}$, pairs of objects included in the positive region are discerned when they have different values of λ_C^β. However, such pairs are not necessarily discerned because there may be a P-reduct such that some of pairs become indiscernible. On the other hand, for each pair u_i and u_j, if they are excluded from the positive region of the common condition attributes, i.e., $C \setminus m_{ij}$, they should be discerned because no subset $A \subseteq C \setminus m_{ij}$ satisfies (VPP1) and (VPP2). From this consideration, discernibility functions characterizing necessary conditions for preservation of the boundaries, the positive region, and the unpredictable region are obtained as follows.

Definition 14 Let $\beta \in [0, 0.5)$ be an admissible error rate. Moreover, let $\tilde{c}_i$ be a Boolean variable pertaining to a condition attribute $c_i \in C$. Discernibility functions $\check{F}_\beta^{\mathrm{B}}$, $\check{F}_\beta^{\mathrm{P}}$, and $\check{F}_\beta^{\mathrm{UN}}$ are defined as follows:

$$\check{F}_\beta^{\mathrm{B}}(\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_m) = \bigwedge_{(i,j) \in \Delta_\beta^{\mathrm{B}}} \bigvee_{c \in m_{ij}} \tilde{c},$$

$$\check{F}_\beta^{\mathrm{P}}(\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_m) = \bigwedge_{(i,j) \in \Delta_\beta^{\mathrm{P}}} \bigvee_{c \in m_{ij}} \tilde{c},$$

$$\check{F}_\beta^{\mathrm{UN}}(\tilde{c}_1, \tilde{c}_2, \ldots, \tilde{c}_m) = \bigwedge_{(i,j) \in \Delta_\beta^{\mathrm{UN}}} \bigvee_{c \in m_{ij}} \tilde{c},$$

where,

$$(i, j) \in \Delta_\beta^{\mathrm{B}} \Leftrightarrow \begin{cases} (\upsilon \backslash \lambda)_C^\beta(u_i) \neq (\upsilon \backslash \lambda)_C^\beta(u_j), \text{ or} \\ (\upsilon \backslash \lambda)_C^\beta(u_i) = (\upsilon \backslash \lambda)_C^\beta(u_j) = \emptyset \\ \quad \text{and } (\upsilon \backslash \lambda)_{C \backslash m_{ij}}^\beta(u_i) = (\upsilon \backslash \lambda)_{C \backslash m_{ij}}^\beta(u_j) \neq \emptyset, \end{cases}$$

$$(i, j) \in \Delta_\beta^{\mathrm{P}} \Leftrightarrow \begin{cases} \lambda_C^\beta(u_i) \neq \emptyset \text{ and } \lambda_C^\beta(u_j) = \emptyset, \text{ or} \\ \lambda_C^\beta(u_i) = \emptyset \text{ and } \lambda_C^\beta(u_j) \neq \emptyset, \text{ or} \\ \lambda_C^\beta(u_i) \neq \emptyset, \ \lambda_C^\beta(u_j) \neq \emptyset, \text{ and } \lambda_{C \backslash m_{ij}}^\beta(u_i) = \lambda_{C \backslash m_{ij}}^\beta(u_j) = \emptyset, \end{cases}$$

$$(i, j) \in \Delta_\beta^{\mathrm{UN}} \Leftrightarrow \begin{cases} \upsilon_C^\beta(u_i) \neq \emptyset \text{ and } \upsilon_C^\beta(u_j) = \emptyset, \text{ or} \\ \upsilon_C^\beta(u_i) = \emptyset \text{ and } \upsilon_C^\beta(u_j) \neq \emptyset, \text{ or} \\ \upsilon_C^\beta(u_i) \neq \emptyset, \ \upsilon_C^\beta(u_j) \neq \emptyset, \text{ and } \upsilon_{C \backslash m_{ij}}^\beta(u_i) = \upsilon_{C \backslash m_{ij}}^\beta(u_j) = \emptyset. \end{cases}$$

Proposition 3 *Let A be a subset of C, and $\beta \in [0, 0.5)$ be an admissible error rate. We have the following implications:*

- *If $\check{F}_\beta^{\mathrm{B}}(\tilde{c}^A) = 0$ then A does not satisfy* (VPB1) *or* (VPB2) *with β,*
- *If $\check{F}_\beta^{\mathrm{P}}(\tilde{c}^A) = 0$ then A does not satisfy* (VPP1) *or* (VPP2) *with β,*
- *If $\check{F}_\beta^{\mathrm{UN}}(\tilde{c}^A) = 0$ then A does not satisfy* (VPUN1) *or* (VPUN2) *with β.*

From Proposition 3, we know that any prime implicant of each of $\check{F}_\beta^{\mathrm{B}}$, $\check{F}_\beta^{\mathrm{P}}$, and $\check{F}_\beta^{\mathrm{UN}}$ can be a subset of some reduct of the corresponding type.

Combining Propositions 2 and 3, we have the following theorem.

Theorem 6 *Let A be a subset of C, and $\beta \in [0, 0.5)$ be an admissible error rate. Let $\hat{\mathscr{P}}_\beta^{\mathrm{B}}$, $\hat{\mathscr{P}}_\beta^{\mathrm{P}}$ and $\hat{\mathscr{P}}_\beta^{\mathrm{UN}}$ be the sets of condition attribute subsets corresponding to the prime implicants of $\hat{F}_\beta^{\mathrm{B}}$, $\hat{F}_\beta^{\mathrm{P}}$, and $\hat{F}_\beta^{\mathrm{UN}}$, respectively. Moreover, let $\check{\mathscr{P}}_\beta^{\mathrm{B}}$, $\check{\mathscr{P}}_\beta^{\mathrm{P}}$ and $\check{\mathscr{P}}_\beta^{\mathrm{UN}}$ be the sets of condition attribute subsets corresponding to the prime implicants of $\check{F}_\beta^{\mathrm{B}}$, $\check{F}_\beta^{\mathrm{P}}$, and $\check{F}_\beta^{\mathrm{UN}}$, respectively. Then, we have the following implications:*

- *If A is a B-reduct with β then $A \in \{B \subseteq C \mid B \supseteq B' \text{ for some } B' \in \check{\mathscr{P}}_\beta^{\mathrm{B}} \text{ and } B \not\supseteq B'' \text{ for any } B'' \in \hat{\mathscr{P}}_\beta^{\mathrm{B}}\}$,*
- *If A is a P-reduct with β then $A \in \{B \subseteq C \mid B \supseteq B' \text{ for some } B' \in \check{\mathscr{P}}_\beta^{\mathrm{P}} \text{ and } B \not\supseteq B'' \text{ for any } B'' \in \hat{\mathscr{P}}_\beta^{\mathrm{P}}\}$,*
- *If A is a UN-reduct with β then $A \in \{B \subseteq C \mid B \supseteq B' \text{ for some } B' \in \check{\mathscr{P}}_\beta^{\mathrm{UN}} \text{ and } B \not\supseteq B'' \text{ for any } B'' \in \hat{\mathscr{P}}_\beta^{\mathrm{UN}}\}$.*

Table 7.4 Discernibility functions related to 9 kinds of reducts

Reduct	Discernibility function(s)	Exact/approximate
L	F_β^{L}	Exact
U	F_β^{U}	Exact
B	$(\check{F}_\beta^{\text{B}}, F_\beta^{\text{L}} \wedge F_\beta^{\text{U}})$	Approximate
P	$(\check{F}_\beta^{\text{P}}, F_\beta^{\text{L}})$	Approximate
UN	$(\check{F}_\beta^{\text{UN}}, F_\beta^{\text{U}})$	Approximate
LU	$F_\beta^{\text{L}} \wedge F_\beta^{\text{U}}$	Exact
PUN	$(\check{F}_\beta^{\text{P}} \wedge \check{F}_\beta^{\text{UN}}, F_\beta^{\text{L}} \wedge F_\beta^{\text{U}})$	Approximate
LUN	$(F_\beta^{\text{L}} \wedge \check{F}_\beta^{\text{UN}}, F_\beta^{\text{L}} \wedge F_\beta^{\text{U}})$	Approximate
BUN	$(\check{F}_\beta^{\text{B}} \wedge \check{F}_\beta^{\text{P}}, F_\beta^{\text{L}} \wedge F_\beta^{\text{U}})$	Approximate

The obtained discernibility functions are shown in Table 7.4. In the case of approximate discernibility functions, the first function in the parenthesis characterizes the necessary condition of the preservation and the second function characterizes the sufficient condition. The discernibility functions related to LU-reducts, LUN-reducts and BUN-reducts can be obtained by taking the conjunctions of discernibility functions related to L-reducts, U-reducts, B-reducts and UN-reducts. Note that $\hat{F}_\beta^{\text{B}} \wedge \hat{F}_\beta^{\text{UN}} = (F_\beta^{\text{L}} \wedge F_\beta^{\text{U}}) \wedge F_\beta^{\text{U}} = F_\beta^{\text{L}} \wedge F_\beta^{\text{U}}$. This is why we have $F_\beta^{\text{L}} \wedge F_\beta^{\text{U}}$ as the discernibility function characterizing a sufficient condition for the preservation of BUN-reducts.

Example 9 Remember the decision table $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ in Table 7.3. Let an admissible error rate be $\beta = 0.39$. In Table 7.5, we show the decision table with three generalized decision functions $\lambda_C^{0.39}, \upsilon_C^{0.39}, (\upsilon \backslash \lambda)_C^{0.39}$ with respect to C and $\beta = 0.39$.

Now let us enumerate reducts as prime implicants of discernibility functions. First let us discuss L-, U- and LU-reducts with $\beta = 0.39$. The discernibility matrix of the decision table is shown as below.

$$
\begin{array}{c|ccccc}
 & P_1 & P_2 & P_3 & P_4 & P_5 \\
\hline
P_1 & \emptyset & \{c_3, c_4\} & \{c_1\} & \{c_1, c_2\} & \{c_2, c_3\} \\
P_2 & \{c_3, c_4\} & \emptyset & \{c_1, c_3, c_4\} & C & \{c_2, c_4\} \\
P_3 & \{c_1\} & \{c_1, c_3, c_4\} & \emptyset & \{c_2\} & \{c_1, c_2, c_3\} \\
P_4 & \{c_1, c_2\} & C & \{c_2\} & \emptyset & \{c_1, c_3\} \\
P_5 & \{c_2, c_3\} & \{c_2, c_4\} & \{c_1, c_2, c_3\} & \{c_1, c_3\} & \emptyset
\end{array}
$$

Table 7.5 The decision table in Table 7.3 with the generalized decision functions

	c_1	c_2	c_3	c_4	d:(b,m,g)	λ_C^β	$\upsilon_C^{0.39}$	$(\upsilon \backslash \lambda)_C^{0.39}$
P_1	Good	Good	Bad	Good	(0,2,9)	{g}	{g}	Ø
P_2	Good	Good	Good	Bad	(0,19,1)	{m}	{m}	Ø
P_3	Bad	Good	Bad	Good	(1,1,2)	Ø	{g}	{g}
P_4	Bad	Bad	Bad	Good	(0,1,1)	Ø	{m,g}	{m,g}
P_5	Good	Bad	Good	Good	(1,1,1)	Ø	Ø	Ø

From the table, we obtain $F_{0.39}^{L}$ and $F_{0.39}^{U}$ as follows:

$$
\begin{aligned}
F_{0.39}^{L}(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4) &= \bigwedge_{i=1,2, j=3,4,5} \bigvee_{c \in m_{ij}} \tilde{c} \wedge \bigvee_{c \in m_{12}} \tilde{c} \\
&= (\tilde{c}_1) \wedge (\tilde{c}_2 \vee \tilde{c}_3) \wedge (\tilde{c}_2 \vee \tilde{c}_4) \wedge (\tilde{c}_3 \vee \tilde{c}_4) \\
&= (\tilde{c}_1 \wedge \tilde{c}_2 \wedge \tilde{c}_3) \vee (\tilde{c}_1 \wedge \tilde{c}_2 \wedge \tilde{c}_4) \vee (\tilde{c}_1 \wedge \tilde{c}_3 \wedge \tilde{c}_4), \\
F_{0.39}^{U}(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4) &= \bigwedge_{(i,j) \in \{(k,l)|k \neq l\} \setminus \{(1,3),(3,1)\}} \bigvee_{c \in m_{ij}} \tilde{c} \\
&= (\tilde{c}_2) \wedge (\tilde{c}_1 \vee \tilde{c}_3) \wedge (\tilde{c}_3 \vee \tilde{c}_4) = (\tilde{c}_2 \wedge \tilde{c}_3) \vee (\tilde{c}_1 \wedge \tilde{c}_2 \wedge \tilde{c}_4).
\end{aligned}
$$

We obtain $F_{0.39}^{LU}$ as:

$$
\begin{aligned}
F_{0.39}^{LU}(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4) &= F^{L}(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4) \wedge F^{U}(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4) \\
&= (\tilde{c}_1) \wedge (\tilde{c}_2) \wedge (\tilde{c}_3 \vee \tilde{c}_4) = (\tilde{c}_1 \wedge \tilde{c}_2 \wedge \tilde{c}_3) \vee (\tilde{c}_1 \wedge \tilde{c}_2 \wedge \tilde{c}_4).
\end{aligned}
$$

Therefore, L-reducts are obtained as $\{c_1, c_2, c_3\}$, $\{c_1, c_2, c_4\}$ and $\{c_1, c_3, c_4\}$. U-reducts are obtained as $\{c_2, c_3\}$ and $\{c_1, c_2, c_4\}$. LU-reducts are obtained as $\{c_1, c_2, c_3\}$ and $\{c_1, c_2, c_4\}$. Note that $\{c_2, c_3\}$ is not an L-reduct but a U-reduct. This is very different from the relation between L- and U-reducts in the classical RSM, i.e., in the classical RSM, a U-reduct includes an L-reduct but an L-reduct never includes a U-reduct.

Now let us discuss B-, P-, UN-reducts with $\beta = 0.39$. We can obtain only approximations of those reducts. To this end, let us get discernibility functions $\check{F}_{0.39}^{B}$, $\check{F}_{0.39}^{P}$, and $\check{F}_{0.39}^{UN}$. For B-reducts, considering the second condition of $\Delta_{0.39}^{B}$, check each pair P_i and P_j such that $(\upsilon \backslash \lambda)_{C}^{0.39}(P_i) = \emptyset$ and $(\upsilon \backslash \lambda)_{C}^{0.39}(P_j) = \emptyset$.

$$
\begin{aligned}
(\upsilon \backslash \lambda)_{C \backslash m_{12}}^{0.39}(P_1) = (\upsilon \backslash \lambda)_{C \backslash m_{12}}^{0.39}(P_2) = \emptyset, \ (\upsilon \backslash \lambda)_{C \backslash m_{15}}^{0.39}(P_1) = (\upsilon \backslash \lambda)_{C \backslash m_{15}}^{0.39}(P_5) = \emptyset, \\
(\upsilon \backslash \lambda)_{C \backslash m_{25}}^{0.39}(P_2) = (\upsilon \backslash \lambda)_{C \backslash m_{25}}^{0.39}(P_5) = \emptyset.
\end{aligned}
$$

For P-reducts, check each pair such that $\lambda_{C}^{0.39}(P_i) \neq \emptyset$ and $\lambda_{C}^{0.39}(P_j) \neq \emptyset$.

$$
\lambda_{C \backslash m_{12}}^{0.39}(P_1) = \lambda_{C \backslash m_{12}}^{0.39}(P_2) = \{\text{medium}\}.
$$

Finally, for UN-reducts, check each pair such that $\upsilon_{C}^{0.39}(P_i) \neq \emptyset$ and $\upsilon_{C}^{0.39}(P_j) \neq \emptyset$.

$$
\begin{aligned}
\upsilon_{C \backslash m_{12}}^{0.39}(P_1) &= \upsilon_{C \backslash m_{12}}^{0.39}(P_2) = \{\text{medium}\}, \quad \upsilon_{C \backslash m_{13}}^{0.39}(P_1) = \upsilon_{C \backslash m_{13}}^{0.39}(P_3) = \{\text{good}\}, \\
\upsilon_{C \backslash m_{14}}^{0.39}(P_1) &= \upsilon_{C \backslash m_{14}}^{0.39}(P_4) = \{\text{good}\}, \quad \upsilon_{C \backslash m_{23}}^{0.39}(P_2) = \upsilon_{C \backslash m_{23}}^{0.39}(P_3) = \{\text{medium}\}, \\
\upsilon_{C \backslash m_{24}}^{0.39}(P_2) &= \upsilon_{C \backslash m_{24}}^{0.39}(P_4) = \{\text{medium}\}, \quad \upsilon_{C \backslash m_{34}}^{0.39}(P_3) = \upsilon_{C \backslash m_{34}}^{0.39}(P_4) = \{\text{good}\}.
\end{aligned}
$$

Therefore, discernibility functions $\check{F}_{0.39}^{B}$, $\check{F}_{0.39}^{P}$, and $\check{F}_{0.39}^{UN}$ are obtained as:

$$\check{F}_{0.39}^{B}(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4) = \bigwedge_{i=1,2,5,\, j=3,4} \bigvee_{c \in m_{ij}} \tilde{c} \wedge \bigvee_{c \in m_{34}} \tilde{c} = \tilde{c}_1 \wedge \tilde{c}_2,$$

$$\check{F}_{0.39}^{P}(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4) = \bigwedge_{i=1,2,\, j=3,4,5} \bigvee_{c \in m_{ij}} \tilde{c} = (\tilde{c}_1) \wedge (\tilde{c}_2 \vee \tilde{c}_3) \wedge (\tilde{c}_2 \vee \tilde{c}_4)$$

$$= (\tilde{c}_1 \wedge \tilde{c}_2) \vee (\tilde{c}_1 \wedge \tilde{c}_3 \wedge \tilde{c}_4),$$

$$\check{F}_{0.39}^{UN}(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \tilde{c}_4) = \bigwedge_{i=1,2,3,4,\, j=5} \bigvee_{c \in m_{ij}} \tilde{c} = (\tilde{c}_1 \vee \tilde{c}_3) \wedge (\tilde{c}_2 \vee \tilde{c}_3) \wedge (\tilde{c}_2 \vee \tilde{c}_4)$$

$$= (\tilde{c}_1 \wedge \tilde{c}_2) \vee (\tilde{c}_2 \wedge \tilde{c}_3) \vee (\tilde{c}_3 \wedge \tilde{c}_4).$$

Because $\hat{F}_{0.39}^{B} = F_{0.39}^{L} \wedge F_{0.39}^{U} = (\tilde{c}_1 \wedge \tilde{c}_2 \wedge \tilde{c}_3) \vee \tilde{c}_1 \wedge \tilde{c}_2 \wedge \tilde{c}_4)$, the candidates of B-reducts are,

$$\{c_1, c_2\}, \quad \{c_1, c_2, c_3\}, \quad \{c_1, c_2, c_4\}.$$

We can see that all of those satisfy (VPB1), hence, $\{c_1, c_2\}$ is the unique B-reduct. Because $\hat{F}_{0.39}^{P} = F_{0.39}^{L} = (\tilde{c}_1 \wedge \tilde{c}_2 \wedge \tilde{c}_3) \vee (\tilde{c}_1 \wedge \tilde{c}_2 \wedge \tilde{c}_4) \vee (\tilde{c}_1 \wedge \tilde{c}_3 \wedge \tilde{c}_4)$, the candidates of P-reducts are,

$$\{c_1, c_2\}, \quad \{c_1, c_2, c_3\}, \quad \{c_1, c_2, c_4\}, \quad \{c_1, c_3, c_4\}.$$

Also, in that case, all candidates satisfy (VPP1), hence, $\{c_1, c_2\}$ and $\{c_1, c_3, c_4\}$ are P-reducts. Similarly, the candidates of UN-reducts are,

$$\{c_1, c_2\}, \quad \{c_2, c_3\}, \quad \{c_3, c_4\}, \quad \{c_1, c_2, c_4\}, \quad \{c_1, c_3, c_4\},$$

and all candidates satisfy (VPUN1), hence, $\{c_1, c_2\}$, $\{c_2, c_3\}$, and $\{c_3, c_4\}$ are UN-reducts.

All reducts are arranged in Table 7.6. We can observe that several kinds of reducts are different. In this example, each L-reduct is also an LUN-reduct and vice versa. Such an equivalence holds in this example but not always.

In this example, we would select $\{c_1, c_2, c_3\}$ or $\{c_1, c_2, c_4\}$ to preserve all structures. Additionally, c_1 and c_2 appear in many other reducts. Whereas, we would select U-reduct $\{c_2, c_3\}$ to reduce the size of the reduct.

Table 7.6 All obtained reducts with $\beta = 0.39$ in Table 7.3

Type	Reducts
L-reduct	$\{c_1, c_2, c_3\}, \{c_1, c_2, c_4\}, \{c_1, c_3, c_4\}$
U-reduct	$\{c_2, c_3\}, \{c_1, c_2, c_4\}$
LU-reduct	$\{c_1, c_2, c_3\}, \{c_1, c_2, c_4\}$
B-reduct	$\{c_1, c_2\}$
P-reduct	$\{c_1, c_2\}, \{c_1, c_3, c_4\}$
UN-reduct	$\{c_1, c_2\}, \{c_2, c_3\}, \{c_3, c_4\}$
LUN-reduct	$\{c_1, c_2, c_3\}, \{c_1, c_2, c_4\}, \{c_1, c_3, c_4\}$
BUN-reduct	$\{c_1, c_2\}$
PUN-reduct	$\{c_1, c_2\}, \{c_1, c_3, c_4\}$

7.4 Structure-Based Attribute Reduction in Dominance-Based Rough Set Models

7.4.1 Decision Tables Under Dominance Principle and Dominance-Based Rough Set Models

In Dominance-based Rough Set Model (DRSM), known as Dominance-based Rough Set Approach [16, 18, 49], decision tables with order relations are analyzed. Let $\mathbb{D} = (U, AT = C \cup \{d\}, \{V_a\}_{a \in AT})$ be a decision table. The attribute set AT is partitioned into AT_N and AT_C, where AT_N is the set of nominal attributes and AT_C is the set of criteria (ordinal attributes). For a criterion $a \in AT_C$, we suppose a total order $\geq$ on its value set V_a. Moreover, all criteria are of the gain-type, i.e., the greater the better. We assume that the decision attribute d is a criterion.

In DRSM, it is supposed that if an object u is better than or equal to another object u' with respect to all condition attributes, then the class of u should not be worse than that of u'. This is called the dominance principle [16].

Remark 8 The setting of DRSM is considered as the monotone or ordinal classification problem [2, 3, 32], where classifiers are restricted to be monotonic. Let f be a classifier, which assigns to each object u a class label (decision class value) $f(u)$. The classifier f is monotonic if for any object pair u and u', we have $u \leq u'$ implying $f(u) \leq f(u')$. In this chapter, however, we do not discuss classifiers nor algorithms for building classifiers.

Remark 9 We assume the total order, i.e., antisymmetry, transitivity, and, comparability, on the value set V_a of each condition criteria $a \in AT_C \cap C$. However, regardless of comparability, the result of this section can be applied without modification. Additionally, we assume that all criteria are of the gain-type. However, in applications, we may encounter cost-type criteria, i.e., the smaller the better. For a cost-type criterion, we can deal with it as the gain-type by reversing the order of its values.

Remark 10 Generally, there is more than one decision attribute in a decision table. In such a case, the set of decision classes (the partition of objects by the decision attributes) is partially ordered, while it is totally ordered in the case of a single decision attribute. In this section, we focus on the case of a single decision attribute (more generally, the case when the decision classes form a totally ordered set), however, the results of this section could be straightforwardly extended to that of multiple decision attributes.

For $A \subseteq C$, a dominance relation D_A on U is defined by:

$$D_A = \Big\{(u, u') \in U^2 | a(u) \geq a(u'), \forall a \in AT_C \cap A$$
$$\text{and } a(u) = a(u'), \forall a \in AT_N \cap A \Big\}.$$

D_A satisfies reflexivity and transitivity. When $(u, u') \in D_A$, we say that u dominates u' with respect to A. The relation $(u, u') \in D_A$ means "u is better than or equal

to u' with respect to criteria A". For $u \in U$, its dominating set and its dominated set with respect to A are defined, respectively, by:

$$D_A^+(u) = \{u' \in U \mid (u', u) \in D_A\},$$
$$D_A^-(u) = \{u' \in U \mid (u, u') \in D_A\}.$$

The dominating set $D_A^+(u)$ (resp. the dominated set $D_A^-(u)$) is the set of the objects dominating (resp. dominated by) u under A.

Since decision classes are ordered $X_1 < X_2 < \cdots < X_p$, one can define an upward union of decision classes $X_i^\geq$ and a downward union of decision classes $X_i^\leq$ with respect to each class X_i, $i \in V_d$, as follows:

$$X_i^\geq = \bigcup_{j \geq i} X_j, \quad X_i^\leq = \bigcup_{j \leq i} X_j.$$

For convenience, $X_0^\leq = X_{p+1}^\geq = \emptyset$. We have $X_i^\geq = U \setminus X_{i-1}^\leq$.

Example 10 Consider a decision table $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ given in Table 7.7. This table shows student evaluation in a school. The objects are seven students, i.e., $U = \{u_1, u_2, \ldots, u_7\}$. The condition attributes are scores of mathematics (Ma), physics (Ph) and literature (Li), while the decision attribute (d) is a comprehensive evaluation (E). Namely, $C = \{$Ma, Ph, Li$\}$ and $d = $E. We may assume that the better scores in all subjects student takes, the better comprehensive evaluation he/she gets.

Let $A = \{$Ma, Ph$\}$. The dominance relation D_A is described as the following matrix. Symbol $*$ indicates that the corresponding row object u_i and column object u_j is in the dominance relation, i.e., $(u_i, u_j) \in D_A$.

	u_1	u_2	u_3	u_4	u_5	u_6	u_7
u_1	*	*	*	*	*	*	*
u_2	*	*	*	*	*	*	*
u_3			*	*	*	*	*
u_4			*				*
u_5					*	*	*
u_6					*	*	*
u_7							*

Table 7.7 A decision table of student records

Student	Ma	Ph	Li	E
u_1	Good	Good	Good	Good
u_2	Good	Good	Med	Med
u_3	Med	Good	Med	Good
u_4	Bad	Med	Good	Med
u_5	Med	Bad	Med	Bad
u_6	Med	Bad	Bad	Med
u_7	Bad	Bad	Bad	Bad

For each object $u_i \in U$, symbols $*$ in the row of u_i indicate the objects in $D^-(u_i)$, while symbols $*$ in the column indicate the objects in $D^+(u_i)$. For example, $D^-(u_3) = \{u_3, u_4, u_5, u_6, u_7\}$ and $D^+(u_3) = \{u_1, u_2, u_3\}$.

There are three decision classes $X_b = \{u_5, u_7\}$, $X_m = \{u_2, u_4, u_6\}$ and $X_g = \{u_1, u_3\}$ for bad, med and good, respectively. The upward and downward unions of those decision classes are,

$$X_b^{\geq} = U, \quad X_m^{\geq} = \{u_1, u_2, u_3, u_4, u_6\}, \quad X_g^{\geq} = \{u_1, u_3\},$$
$$X_b^{\leq} = \{u_5, u_7\}, \quad X_m^{\leq} = \{u_2, u_4, u_5, u_6, u_7\}, \quad X_g^{\leq} = U.$$

Given a decision table, the inconsistency with respect to the dominance principle is captured by the difference between upper and lower approximations of the unions of decision classes. Given a condition attribute set $A \subseteq C$, and $i \in V_d$, the lower approximation $\text{LA}_A(X_i^{\geq})$ and the upper approximation $\text{UA}_A(X_i^{\geq})$ of $X_i^{\geq}$ are defined, respectively, by:

$$\text{LA}_A(X_i^{\geq}) = \{u \in U \mid D_A^+(u) \subseteq X_i^{\geq}\},$$
$$\text{UA}_A(X_i^{\geq}) = \{u \in U \mid D_A^-(u) \cap X_i^{\geq} \neq \emptyset\}.$$

Similarly, the lower approximation $\text{LA}_A(X_i^{\leq})$ of $X_i^{\leq}$ and upper approximation $\text{UA}_A(X_i^{\leq})$ are defined, respectively, by:

$$\text{LA}_A(X_i^{\leq}) = \{u \in U \mid D_A^-(u) \subseteq X_i^{\leq}\},$$
$$\text{UA}_A(X_i^{\leq}) = \{u \in U \mid D_A^+(u) \cap X_i^{\leq} \neq \emptyset\}.$$

If u belongs to $\text{LA}_A(X_i^{\geq})$ then all objects dominating u do not belong to $X_{i-1}^{\leq}$, i.e., there exists no evidence for $u \in X_{i-1}^{\leq}$ in view of the monotonicity assumption. Therefore, we can say that u certainly belongs to $X_i^{\geq}$. On the other hand if u belongs to $\text{UA}_A(X_i^{\geq})$ then u is dominating an object belonging to $X_i^{\geq}$, i.e., there exists evidence for $u \in X_i^{\geq}$ in view of the monotonicity assumption. Therefore, we can say that u possibly belongs to $X_i^{\geq}$. The similar interpretations can be applied to $\text{LA}_A(X_i^{\leq})$ and $\text{UA}_A(X_i^{\leq})$.

The difference between the upper and lower approximations is called a boundary. The boundaries of an upward union $X_i^{\geq}$ and a downward union $X_i^{\leq}$, denoted by $\text{BN}_A(X_i^{\geq})$ and $\text{BN}_A(X_i^{\leq})$, are defined by:

$$\text{BN}_A(X_i^{\geq}) = \text{UA}_A(X_i^{\geq}) \setminus \text{LA}_A(X_i^{\geq}),$$
$$\text{BN}_A(X_i^{\leq}) = \text{UA}_A(X_i^{\leq}) \setminus \text{LA}_A(X_i^{\leq}).$$

Objects in the boundary region of an upward or downward union are classified neither to that union nor to the complement with certainty.

Example 11 Remember the decision table $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ in Table 7.7. Let $A = \{\text{Ma, Ph}\}$. The lower and upper approximations of the upward and downward unions with respect to A are obtained as follows.

$$\text{LA}_A(X_{\bar{b}}^{\geq}) = U, \quad \text{LA}_A(X_{\bar{m}}^{\geq}) = \{u_1, u_2, u_3, u_4\}, \quad \text{LA}_A(X_{\bar{g}}^{\geq}) = \emptyset,$$

$$\text{LA}_A(X_{\bar{b}}^{\leq}) = \{u_7\}, \quad \text{LA}_A(X_{\bar{m}}^{\leq}) = \{u_4, u_5, u_6, u_7\}, \quad \text{LA}_A(X_{\bar{g}}^{\leq}) = U,$$

$$\text{UA}_A(X_{\bar{b}}^{\geq}) = U, \quad \text{UA}_A(X_{\bar{m}}^{\geq}) = U \setminus \{u_7\}, \quad \text{UA}_A(X_{\bar{g}}^{\geq}) = \{u_1, u_2, u_3\},$$

$$\text{UA}_A(X_{\bar{b}}^{\leq}) = \{u_5, u_6, u_7\}, \quad \text{UA}_A(X_{\bar{m}}^{\leq}) = U, \quad \text{UA}_A(X_{\bar{g}}^{\leq}) = U.$$

Now, we remember properties of approximations [16, 18, 31]. By the boundary conditions of $X^{\geq}$ and $X^{\leq}$,

$$\text{UA}_A(X_1^{\geq}) = \text{LA}_A(X_1^{\geq}) = U, \quad \text{UA}_A(X_p^{\leq}) = \text{LA}_A(X_p^{\leq}) = U,$$
$$\text{UA}_A(X_{p+1}^{\geq}) = \text{LA}_A(X_{p+1}^{\geq}) = \emptyset, \quad \text{UA}_A(X_0^{\leq}) = \text{LA}_A(X_0^{\leq}) = \emptyset. \tag{7.21}$$

Let $A \subseteq C$ and $i \in V_d$. Similarly to RSM, there exist inclusion relations between each union of decision classes and its lower and upper approximations.

$$\text{LA}_A(X_i^{\geq}) \subseteq X_i^{\geq} \subseteq \text{UA}_A(X_i^{\geq}), \quad \text{LA}_A(X_i^{\leq}) \subseteq X_i^{\leq} \subseteq \text{UA}_A(X_i^{\leq}). \tag{7.22}$$

Approximations are expressed by unions of dominating or dominated sets,

$$\text{LA}_A(X_i^{\geq}) = \bigcup_{D_A^+(u) \subseteq X_i^{\geq}} D_A^+(u) = \bigcup_{u \in \text{LA}_A(X_i^{\geq})} D_A^+(u),$$

$$\text{UA}_A(X_i^{\geq}) = \bigcup_{D_A^-(u) \cap X_i^{\geq} \neq \emptyset} D_A^+(u) = \bigcup_{u \in \text{UA}_A(X_i^{\geq})} D_A^+(u),$$

$$\text{LA}_A(X_i^{\leq}) = \bigcup_{D_A^-(u) \subseteq X_i^{\leq}} D_A^-(u) = \bigcup_{u \in \text{LA}_A(X_i^{\leq})} D_A^-(u),$$

$$\text{UA}_A(X_i^{\leq}) = \bigcup_{D_A^+(u) \cap X_i^{\leq} \neq \emptyset} D_A^-(u) = \bigcup_{u \in \text{UA}_A(X_i^{\leq})} D_A^-(u).$$

There exists duality of lower and upper approximations.

$$\text{UA}_A(X_i^{\geq}) = U \setminus \text{LA}_A(X_{i-1}^{\leq}), \quad \text{UA}_A(X_i^{\leq}) = U \setminus \text{LA}_A(X_{i+1}^{\geq}). \tag{7.23}$$

So, the upper approximations of the pair of complementary unions of decision classes form a cover of U:

$$\text{UA}_A(X_i^{\geq}) \cup \text{UA}_A(X_{i-1}^{\leq}) = U. \tag{7.24}$$

By the duality of lower and upper approximations, the boundaries of the pair of complementary unions are the same,

$$\text{BN}_A(X_i^{\geq}) = \text{BN}_A(X_{i-1}^{\leq}). \tag{7.25}$$

Lower and upper approximations can be expressed by boundaries. That is useful for investigating relations between different types of reducts:

$$\text{UA}_A(X_i^{\geq}) = \text{BN}_A(X_i^{\geq}) \cup X_i^{\geq}, \quad \text{UA}_A(X_i^{\leq}) = \text{BN}_A(X_i^{\leq}) \cup X_i^{\leq}, \tag{7.26}$$

$$\text{LA}_A(X_i^{\geq}) = X_i^{\geq} \setminus \text{BN}_A(X_i^{\geq}), \quad \text{LA}_A(X_i^{\leq}) = X_i^{\leq} \setminus \text{BN}_A(X_i^{\leq}). \tag{7.27}$$

Let $A, B \subseteq C$ and $i, j \in V_d$. Then, we have the following monotonicity properties:

$$j \geq i \Rightarrow \text{LA}_A(X_j^{\geq}) \subseteq \text{LA}_A(X_i^{\geq}), \quad \text{UA}_A(X_j^{\geq}) \subseteq \text{UA}_A(X_i^{\geq}), \tag{7.28}$$

$$j \leq i \Rightarrow \text{LA}_A(X_j^{\leq}) \subseteq \text{LA}_A(X_i^{\leq}), \quad \text{UA}_A(X_j^{\leq}) \subseteq \text{UA}_A(X_i^{\leq}), \tag{7.29}$$

$$B \subseteq A \Rightarrow \text{LA}_B(X_i^{\geq}) \subseteq \text{LA}_A(X_i^{\geq}), \quad \text{LA}_B(X_i^{\leq}) \subseteq \text{LA}_A(X_i^{\leq}), \tag{7.30}$$

$$B \subseteq A \Rightarrow \text{UA}_B(X_i^{\geq}) \supseteq \text{UA}_A(X_i^{\geq}), \quad \text{UA}_B(X_i^{\leq}) \supseteq \text{UA}_A(X_i^{\leq}). \tag{7.31}$$

Those are important for defining and enumerating reducts.

Furthermore, the authors proposed lower and upper approximations and boundary regions of decision classes [31]. For $A \subseteq C$ and $i \in V_d$, lower and upper approximations of X_i and the boundary region of X_i are defined by:

$$\text{LA}_A(X_i) = \text{LA}_A(X_i^{\geq}) \cap \text{LA}_A(X_i^{\leq}),$$

$$\text{UA}_A(X_i) = \text{UA}_A(X_i^{\geq}) \cap \text{UA}_A(X_i^{\leq}),$$

$$\text{BN}_A(X_i) = \text{UA}_A(X_i) \setminus \text{LA}_A(X_i).$$

This definition is an analogy to $X_i = X_i^{\geq} \cap X_i^{\leq}$.

Let $A \subseteq C$ and $i \in V_d$. The upper approximations of $X_i^{\geq}$ and $X_i^{\leq}$ are represented by upper approximations of decision classes:

$$\text{UA}_A(X_i^{\geq}) = \bigcup_{j \geq i} \text{UA}_A(X_j), \tag{7.32}$$

$$\text{UA}_A(X_i^{\leq}) = \bigcup_{j \leq i} \text{UA}_A(X_j). \tag{7.33}$$

The boundary of X_i is the union of the boundaries of $X_i^{\geq}$ and $X_i^{\leq}$,

$$\text{BN}_A(X_i) = \text{BN}_A(X_i^{\geq}) \cup \text{BN}_A(X_i^{\leq}). \tag{7.34}$$

Approximations of decision classes have similar properties as those of unions of decision classes:

$$LA_A(X_i) \subseteq X_i \subseteq UA_A(X_i), \tag{7.35}$$

$$UA_A(X_i) = BN_A(X_i) \cup X_i, \tag{7.36}$$

$$LA_A(X_i) = X_i \setminus BN_A(X_i). \tag{7.37}$$

The next properties are analogies to (7.6) and (7.5) of the classical RSM.

$$BN_A(X_i) = UA_A(X_i) \cap \bigcup_{j \neq i} UA_A(X_j), \tag{7.38}$$

$$LA_A(X_i) = U \setminus \bigcup_{j \neq i} UA_A(X_j). \tag{7.39}$$

We define the positive region for the decision table in DRSM:

$$POS_A(d) = \bigcup_{i \in V_d} LA_A(X_i).$$

The complement of the positive region is exactly the union of all boundaries,

$$U \setminus POS_A(d) = \bigcup_{i \in V_d} BN_A(X_i). \tag{7.40}$$

Moreover, the approximations are also monotone with respect to the inclusion relation between condition attribute sets. Let $A, B \subseteq C$ and $i \in V_d$.

$$B \subseteq A \Rightarrow LA_B(X_i) \subseteq LA_A(X_i), \quad UA_B(X_i) \supseteq UA_A(X_i). \tag{7.41}$$

The generalized decision function proposed by Dembczyński et al. [10] also plays an important role for Boolean reasoning in DRSM. It provides an object-wise view of DRSM. Let $A \subseteq C$ and $u \in U$, generalized decision of u with respect to A is defined by $\delta_A(u) = \langle l_A(u), u_A(u) \rangle$, where

$$l_A(u) = \min\{i \in V_d \mid D_A^+(u) \cap X_i \neq \emptyset\},$$
$$u_A(u) = \max\{i \in V_d \mid D_A^-(u) \cap X_i \neq \emptyset\}.$$

$\delta_A(u)$ shows the interval of decision classes to which x may belong. $l_A(u)$ and $u_A(u)$ are the lower and upper bounds of the interval. Obviously, we have

$$l_A(u) \leq u_A(u). \tag{7.42}$$

$l_A(u)$ and $u_A(u)$ are monotone with respect to the inclusion relation between condition attribute sets. Namely, for $B, A \subseteq C$ and $u \in U$, we have

$$B \subseteq A \Rightarrow l_B(u) \le l_A(u), \quad u_B(u) \ge u_A(u). \tag{7.43}$$

Let $i \in V_d$, using the generalized decision function, the lower and upper approximations of unions are represented as:

$$\text{LA}_A(X_i^\ge) = \{u \in U \mid l_A(u) \ge i\}, \quad \text{UA}_A(X_i^\ge) = \{u \in U \mid u_A(u) \ge i\}, \tag{7.44}$$
$$\text{LA}_A(X_i^\le) = \{u \in U \mid u_A(u) \le i\}, \quad \text{UA}_A(X_i^\le) = \{u \in U \mid l_A(u) \le i\}. \tag{7.45}$$

We can represent approximations of classes using the generalized decision,

$$\text{LA}_A(X_i) = \{u \in U \mid l_A(u) = u_A(u) = i\}, \tag{7.46}$$
$$\text{UA}_A(X_i) = \{u \in U \mid l_A(u) \le i \le u_A(u)\}, \tag{7.47}$$
$$\text{BN}_A(X_i) = \{u \in U \mid l_A(u) \le i \le u_A(u), \, l_A(u) < u_A(u)\}. \tag{7.48}$$

Example 12 Remember the decision table $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ in Table 7.7. Let $A = \{\text{Ma}, \text{Ph}\}$. The generalized decision function δ_A with respect to A is obtained as follows:

$\delta_A(u_1) = \langle \text{med}, \text{good} \rangle, \quad \delta_A(u_2) = \langle \text{med}, \text{good} \rangle, \quad \delta_A(u_3) = \langle \text{med}, \text{good} \rangle,$
$\delta_A(u_4) = \langle \text{med}, \text{med} \rangle, \quad \delta_A(u_5) = \langle \text{bad}, \text{med} \rangle, \quad \delta_A(u_6) = \langle \text{bad}, \text{med} \rangle,$
$\delta_A(u_7) = \langle \text{bad}, \text{bad} \rangle.$

7.4.2 Structure-Based Reducts in Dominance-Based Rough Set Models

Before defining structure-based reducts in DRSM, we introduce a notion of reducts preserving the quality of sorting, proposed by Susmaga et al. [49]. For $A \subseteq C$, the quality of sorting $\gamma_A(d)$, which is the counterpart of the quality of classification in the classical RSM, is defined by:

$$\gamma_A(d) = \frac{|U - \bigcup_{i \in V_d} \text{BN}_A(X_i^\le)|}{|U|} = \frac{|U - \bigcup_{i \in V_d} \text{BN}_A(X_i^\ge)|}{|U|}.$$

By (7.34) and (7.40), we can see that $\gamma_A(d)$ is related to the positive region of DRSM,

$$\gamma_A(d) = \frac{|U - \bigcup_{i \in V_d} \text{BN}_A(X_i^\le)|}{|U|} = \frac{|U - \bigcup_{i \in V_d} \text{BN}_A(X_i)|}{|U|} = \frac{|\text{POS}_A(d)|}{|U|}.$$

We call this type of reducts Q-reducts.

Definition 15 ([16, 49]) A Q-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\gamma_A(d) = \gamma_C(d). \tag{DQ}$$

Now, we introduce structure-based reducts in DRSM. Lower and upper approximations and boundary regions of upward and downward unions can be considered as a structure over a given object set U. From this point, we define 7 union-structure-preserving reducts. The following reducts are conceivable.

Definition 16 ([25, 52]) We define 7 types of reducts as follows.

- An $L^{\geq}$-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{LA}_A(X_i^{\geq}) = \mathrm{LA}_C(X_i^{\geq}) \quad \text{for all } i \in V_d. \tag{DL$^{\geq}$}$$

- An $L^{\leq}$-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{LA}_A(X_i^{\leq}) = \mathrm{LA}_C(X_i^{\leq}) \quad \text{for all } i \in V_d. \tag{DL$^{\leq}$}$$

- A $U^{\geq}$-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{UA}_A(X_i^{\geq}) = \mathrm{UA}_C(X_i^{\geq}) \quad \text{for all } i \in V_d. \tag{DU$^{\geq}$}$$

- A $U^{\leq}$-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{UA}_A(X_i^{\leq}) = \mathrm{UA}_C(X_i^{\leq}) \quad \text{for all } i \in V_d. \tag{DU$^{\leq}$}$$

- An $L^{\diamond}$-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{LA}_A(X_i^{\geq}) = \mathrm{LA}_C(X_i^{\geq}) \quad \text{and} \quad \mathrm{LA}_A(X_i^{\leq}) = \mathrm{LA}_C(X_i^{\leq}) \quad \text{for all } i \in V_d. \tag{DL$^{\diamond}$}$$

- A $U^{\diamond}$-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{UA}_A(X_i^{\geq}) = \mathrm{UA}_C(X_i^{\geq}) \quad \text{and} \quad \mathrm{UA}_A(X_i^{\leq}) = \mathrm{UA}_C(X_i^{\leq}) \quad \text{for all } i \in V_d. \tag{DU$^{\diamond}$}$$

- A $B^{\diamond}$-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\begin{aligned}
\mathrm{BN}_A(X_i^{\geq}) &= \mathrm{BN}_C(X_i^{\geq}) \quad \text{for all } i \in V_d, \text{ or equivalently,} \\
\mathrm{BN}_A(X_i^{\leq}) &= \mathrm{BN}_C(X_i^{\leq}) \quad \text{for all } i \in V_d.
\end{aligned} \tag{DB$^\diamond$}$$

Yang et al. [52] independently proposed four kinds of reducts in DRSM with unknown attribute values, which are application of distribution reducts of Mi et al. [33]. Those reducts preserve lower/upper approximations of upward/downward unions. Hence, they correspond to $L^{\geq}$-, $L^{\leq}$-, $U^{\geq}$-, and $U^{\leq}$-reducts of ours. However, Yang et al. did not consider boundaries and combinations of different types of reducts.

From (7.23), we know that (DL$^{\geq}$) and (DU$^{\leq}$) are equivalent. Similarly, (DL$^{\leq}$) and (DU$^{\geq}$) are also equivalent. Therefore, (DL$^{\diamond}$) is equivalent to (DU$^{\diamond}$). Moreover, since condition (DL$^{\diamond}$) implies conditions (DL$^{\geq}$) and (DL$^{\leq}$), any $L^{\diamond}$-reduct satisfies (DL$^{\geq}$) and also (DL$^{\leq}$). Similarly, since condition (DU$^{\diamond}$) implies conditions (DU$^{\geq}$) and (DU$^{\leq}$), any $U^{\diamond}$-reduct satisfies (DU$^{\geq}$) and also (DU$^{\leq}$). Therefore, we have the following theorem.

Theorem 7 ([25, 52]) *Let A be a subset of C. The following statements hold.*

- *A is a $U^{\geq}$-reduct if and only if A is an $L^{\leq}$-reduct.*
- *A is a $U^{\leq}$-reduct if and only if A is an $L^{\geq}$-reduct.*
- *A is a $U^{\diamond}$-reduct if and only if A is an $L^{\diamond}$-reduct.*
- *A is a $B^{\diamond}$-reduct if and only if A is an $L^{\diamond}$-reduct.*
- *If A is an $L^{\diamond}$-reduct then A satisfies (DL$^{\geq}$), (DL$^{\leq}$), (DU$^{\geq}$), and (DU$^{\leq}$).*

As the result of the discussion, we obtain 3 different types of reducts based on the structure induced from rough set operations on unions. They are represented by $L^{\geq}$-reduct, $L^{\leq}$-reduct and $L^{\diamond}$-reduct.

Now, we are ready to define other types of structure-based reducts, considering approximations of decision classes. The first kind of reducts, called L-reduct, preserves the lower approximations of decision classes, the second kind of reducts, called U-reduct, preserves the upper approximations of decision classes, the third kind of reduct, called B-reduct, preserves the boundary regions of decision classes, and the fourth kind of reduct, called P-reduct, preserves the positive region. They are parallel to L-, U-, B-, P-reducts discussion in the classical RSM.

Definition 17 ([31]) We define four types of reducts as follows.

- An L-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{LA}_A(X_i) = \mathrm{LA}_C(X_i) \quad \text{for all } i \in V_d. \tag{DL}$$

- A U-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{UA}_A(X_i) = \mathrm{UA}_C(X_i) \quad \text{for all } i \in V_d. \tag{DU}$$

- A B-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{BN}_A(X_i) = \mathrm{BN}_C(X_i) \quad \text{for all } i \in V_d. \tag{DB}$$

- A P-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{POS}_A(d) = \mathrm{POS}_C(d). \tag{DP}$$

From the properties of approximations of decision classes, we have the following theorem.

Theorem 8 ([31]) *Let A be a subset of C. We have the following assertions:*

(a) *A is a B-reduct if and only if A is a U-reduct,*
(b) *A is a P-reduct if and only if A is an L-reduct,*
(c) *If A is a U-reduct then A satisfies (DL).*

Consequently, we have only 2 kinds of class-structure-based reducts: L-reducts and U-reducts (or B-reducts). This result is also parallel to the result in RSM.

Let us discuss relations of the union-based reducts, the class-based reducts, the Q-reducts. We have the following theorems.

Theorem 9 ([31]) *Let A be a subset of C. We have the following assertions:*

(a) *A is an $L^\diamond$-reduct if and only if A is a U-reduct,*
(b) *A is a Q-reduct if and only if A is an L-reduct.*

Additionally, we propose two more types of reducts, which are compounds of L- with $L^\geq$- and $L^\leq$-reducts, respectively.

Definition 18 We define two types of reducts as follows.

- An $LL^\geq$-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:
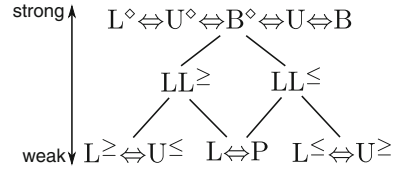
$$\mathrm{LA}_A(X_i) = \mathrm{LA}_C(X_i) \text{ and } \mathrm{LA}_A(X_i^\geq) = \mathrm{LA}_C(X_i^\geq) \quad \text{for all } i \in V_d. \ (\mathrm{DLL}^\geq)$$

- A $LL^\leq$-reduct in DRSM is a minimal condition attribute subset $A \subseteq C$ satisfying the following condition:

$$\mathrm{UA}_A(X_i) = \mathrm{UA}_C(X_i) \text{ and } \mathrm{LA}_A(X_i^\leq) = \mathrm{LA}_C(X_i^\leq) \quad \text{for all } i \in V_d. \ (\mathrm{DLL}^\leq)$$

As a result, all types of reducts proposed in DRSM are arranged in Fig. 7.3. Consequently, there exist six different kinds of reducts, i.e., U-reducts (B-reducts, $L^\diamond$-reducts, $U^\diamond$-reducts, $B^\diamond$-reducts), $LL^\geq$-reducts, $LL^\leq$-reducts, L-reducts (P-reducts), $L^\geq$-reducts ($U^\leq$-reducts) and $L^\leq$-reducts ($U^\geq$-reducts) in DRSM.

Fig. 7.3 Strong-weak hierarchy of reducts in DRSM

$$\text{strong} \uparrow \quad L^{\diamond} \Leftrightarrow U^{\diamond} \Leftrightarrow B^{\diamond} \Leftrightarrow U \Leftrightarrow B$$

$$LL^{\geq} \qquad LL^{\leq}$$

$$\text{weak} \downarrow \ L^{\geq} \Leftrightarrow U^{\leq} \quad L \Leftrightarrow P \quad L^{\leq} \Leftrightarrow U^{\geq}$$

7.4.3 Boolean Functions Representing Reducts

Because Boolean reasoning is a popular approach to enumerate all reducts of each type in rough set literature, some authors already showed Boolean functions representing their own types of reducts [49, 52]. On the other hand, the authors proposed a unified formulation of Boolean functions for all types of reducts using the generalized decision function in [31]. We only discuss Boolean functions for $L^{\geq}$-, $L^{\leq}$- and L-reducts, because U-reducts, $LL^{\geq}$-reducts, $LL^{\leq}$-reducts, and their equivalences can be computed from $L^{\geq}$- and $L^{\leq}$-reducts or their Boolean functions.

 We represent preserving conditions of reducts by those of the generalized decision function.

Lemma 5 ([31]) *Let A be a subset of C. We have the following assertions.*

- *Condition ($DL^{\geq}$) is equivalent to:*

$$l_A(u) = l_C(u) \quad \text{for all } u \in U. \tag{DlG}$$

- *Condition ($DL^{\leq}$) is equivalent to:*

$$u_A(u) = u_C(u) \quad \text{for all } u \in U. \tag{DuG}$$

- *Condition (DL) is equivalent to:*

$$\delta_A(u) = \delta_C(u) \quad \text{for all } u \in U \text{ such that } l_C(u) = u_C(u). \tag{DLG}$$

 The next lemma is parallel to Lemma 2 of RSM. It also connects two notions: "preserving" and "non-dominating".

Lemma 6 ([31]) *Let $u \in U$. The following assertions are equivalent.*

- $l_A(u) = l_C(u).$
- $\forall u' \in U, \ (l_C(u') < l_C(u) \Rightarrow \exists a \in A, \ (u', u) \notin D_{\{a\}}).$

Moreover, the following assertions are also equivalent.

- $u_A(u) = u_C(u).$
- $\forall u' \in U, \ (u_C(u') > u_C(u) \Rightarrow \exists a \in A, \ (u, u') \notin D_{\{a\}}).$

Now we are ready to define a non-domination matrix, instead of the discernibility matrix of RSM. The non-domination matrix $M = (m_{ij})_{i,j=1,2,...,n}$ in DRSM is defined as follows:

$$m_{ij} = \{c \in C \mid (u_j, u_i) \notin D_{\{c\}}\}$$

Based on M, we define four non-domination functions.

Definition 19 Non-domination functions $F^{\geq}$, $F^{\leq}$ and F^{L} are defined as follows.

$$F^{\geq}(\tilde{c}_1, \ldots, \tilde{c}_m) = \bigwedge_{i,j \mid l_C(u_j) < l_C(u_i)} \bigvee_{c \in m_{ij}} \tilde{c},$$

$$F^{\leq}(\tilde{c}_1, \ldots, \tilde{c}_m) = \bigwedge_{i,j \mid u_C(u_j) > u_C(u_i)} \bigvee_{c \in m_{ji}} \tilde{c},$$

$$F^{L}(\tilde{c}_1, \ldots, \tilde{c}_m) = \bigwedge_{i:l_C(x_i) = u_C(x_i)} \left(\bigwedge_{j \mid l_C(u_j) < l_C(u_i)} \bigvee_{c \in m_{ij}} \tilde{c} \wedge \bigwedge_{j \mid u_C(u_j) > u_C(u_i)} \bigvee_{c \in m_{ji}} \tilde{c} \right),$$

where $\tilde{c}_i$ is a Boolean variable corresponding to ith condition attribute c_i.

From Lemma 6, we have the following theorem. Let $A \subseteq C$. Remember that $\tilde{c}^A$ is a Boolean vector such that ith element $\tilde{c}_i^A$ is true iff $c_i \in A$, and ϕ_A is the term $\bigwedge\{\tilde{c} \mid c \in A\}$.

Theorem 10 ([31, 49, 52]) *Let A be a subset of C. We have the following equivalences:*

- *A satisfies (DlG), i.e., (DL$^{\geq}$) if and only if $F^{\geq}(\tilde{c}^A) = 1$. Moreover, A is an $L^{\geq}$-reduct in DRSM if and only if ϕ_A is a prime implicant of $F^{\geq}$,*
- *A satisfies (DuG), i.e., (DL$^{\leq}$) if and only if $F^{\leq}(\tilde{c}^A) = 1$. Moreover, A is an $L^{\leq}$-reduct in DRSM if and only if ϕ_A is a prime implicant of $F^{\leq}$,*
- *A satisfies (DLG), i.e., (DL) if and only if $F^{L}(\tilde{c}^A) = 1$. Moreover, A is an L-reduct in DRSM if and only if ϕ_A is a prime implicant of F^{L}.*

From Theorem 10, all $L^{\geq}$-, $L^{\leq}$- and L-reducts can be obtained as all prime implicants of Boolean functions $F^{\geq}$, $F^{\leq}$ and F^{L}, respectively.

The proposed non-domination matrices have an advantage when compared with the previous ones. We need to calculate neither lower, upper approximations nor boundary regions of unions but only the lower bounds l_C and the upper bounds u_C of all objects. Namely, the computation of the proposed approach is free from the number of decision classes.

Example 13 Remember the decision table given $\mathbb{D} = (U, C \cup \{d\}, \{V_a\})$ in Table 7.7. In Table 7.8, we show again $\mathbb{D}$ with the lower bounds l_C and the upper bounds u_C of the generalized decisions of the objects which appear in the rightmost two columns of the table. To obtain $l_C(u_i)$ and $u_C(u_i)$, we search the minimum class in $D_C^+(u_i)$ and the maximum class in $D_C^-(u_i)$, respectively.

Table 7.8 The decision table in Table 7.7 with the generalized decision function

Student	Ma	Ph	Li	E	$\delta_C = \langle l_C, u_C \rangle$
u_1	Good	Good	Good	Good	$\langle good, good \rangle$
u_2	Good	Good	Med	Med	$\langle med, good \rangle$
u_3	Med	Good	Med	Good	$\langle med, good \rangle$
u_4	Bad	Med	Good	Med	$\langle med, med \rangle$
u_5	Med	Bad	Med	Bad	$\langle bad, med \rangle$
u_6	Med	Bad	Bad	Med	$\langle bad, med \rangle$
u_7	Bad	Bad	Bad	Bad	$\langle bad, bad \rangle$

Non-domination matrices M are obtained as follows.

	u_1	u_2	u_3	u_4	u_5	u_6	u_7
u_1^*	∅	{Li}	{Ma, Li}	{Ma, Ph}	C	C	C
u_2	∅	∅	{Ma}	{Ma, Ph}	{Ma, Ph}	C	C
u_3	∅	∅	∅	{Ma, Ph}	{Ph}	{Ph, Li}	C
u_4^*	∅	{Li}	{Li}	∅	{Ph, Li}	{Ph, Li}	{Ph, Li}
u_5	∅	∅	∅	{Ma}	∅	{Li}	{Ma, Li}
u_6	∅	∅	∅	{Ma}	∅	∅	{Ma}
u_7^*	∅	∅	∅	∅	∅	∅	∅

For example, the entry corresponding to row u_1 and column u_3 on M contains Ma and Li, because u_3 is worse than u_1 with respect to Ma and Li but not worse with respect to Ph. Symbol C at some entries means {Ma, Ph, Li}. The rows with symbol $*$ show objects u_i such that $l_C(u_i) = u_C(u_i)$.

The Boolean function $F^\geq$ is obtained from M as

$$F^\geq(\tilde{Ma}, \tilde{Ph}, \tilde{Li}) = \bigwedge_{i=1, \ j=2,3,\ldots,7} \bigvee_{c \in m_{ij}} \tilde{c} \wedge \bigwedge_{i=2,3,4, \ j=5,6,7} \bigvee_{c \in m_{ij}} \tilde{c} = \tilde{Ph} \wedge \tilde{Li}.$$

From the last equation, $F^\geq(\tilde{Ma}, \tilde{Ph}, \tilde{Li}) =$ true only when $\tilde{Ph} =$ true and $\tilde{Li} =$ true. This implies that only {Ma, Ph, Li} and {Ph, Li} satisfy (DL$^\geq$) owing to Theorem 10. An L$^\geq$-reduct is a minimal set of condition attributes that satisfies (DL$^\geq$). Therefore, {Ph, Li} is a unique L$^\geq$-reduct. Moreover, the L$^\geq$-reduct corresponds to a unique prime implicant of $F^\geq$, i.e., $\tilde{Ph} \wedge \tilde{Li}$.

Similarly, Boolean functions $F^\leq$, F^U and F^L are

$$F^\leq(\tilde{Ma}, \tilde{Ph}, \tilde{Li}) = \bigwedge_{i=4,5,6,7, \ j=1,2,3} \bigvee_{c \in m_{ji}} \tilde{c} \wedge \bigwedge_{i=7, \ j=1,2,\ldots,6} \bigvee_{c \in m_{ji}} \tilde{c} = \tilde{Ma} \wedge \tilde{Ph},$$

$$F^L(\tilde{Ma}, \tilde{Ph}, \tilde{Li}) = \bigwedge_{i=1, \ j=2,3,\ldots,7} \bigvee_{c \in m_{ij}} \tilde{c} \wedge \bigwedge_{i=4, \ j=5,6,7} \bigvee_{c \in m_{ij}} \tilde{c} \wedge \bigwedge_{i=4, \ j=1,2,3} \bigvee_{c \in m_{ji}} \tilde{c}$$
$$\wedge \bigwedge_{i=7 \ j=1,2,\ldots,6} \bigvee_{c \in m_{ji}} \tilde{c} = \tilde{Ma} \wedge \tilde{Li}.$$

Consequently, we obtain {Ph, Li} as the unique $L^{\geq}$-reduct, {Ma, Ph} as the unique $L^{\leq}$-reduct and {Ma, Li} as the unique L-reduct. Moreover, {Ph, Li} ∪ {Ma, Ph} = {Ma, Ph, Li} = C is the unique U-reduct, {Ph, Li} ∪ {Ma, Li} = {Ma, Ph, Li} = C is the unique $LL^{\geq}$-reduct, and {Ma, Ph} ∪ {Ma, Li} = {Ma, Ph, Li} = C is the unique $LL^{\leq}$-reduct.

7.5 Concluding Remarks

In this chapter, we have studied structure-based attribute reduction as a rough set approach to the attribute selection/reduction problem. We have proposed several concepts of structure-based reducts. In the rough set model, there are 2 different types of reducts, U-reducts and L-reducts. U-reducts preserve generalized decisions $\partial_C(u)$ for all objects $u \in U$, while L-reducts do so for all certain classified objects u, namely, $|\partial_C(u)| = 1$. The authors studied refinement of the hierarchy of structure-based reducts (Fig. 7.1) by interpolating reducts which preserve objects u whose generalized decisions are at most k, namely, $|\partial_C(u)| \leq k$ [24]. The parameter k provides a trade-off between the size of a reduct and preserved information.

In VPRSM, because approximations may not be monotone with respect to the set inclusion of condition attributes, classifications of some objects become precise by reducing condition attributes. From that viewpoint, the authors have proposed enhancing reducts [21], which do not preserve but make classification more precise than that of all condition attributes.

Attribute reduction have been also studied in other extensions of the rough set model, e.g. tolerance-based RSM [44], RSM for decision tables with missing values [29, 30], Bayesian RSM [47], fuzzy RSM [27, 28], and variable precision DRSM [26]. However, in general, extensions of the rough set model drop some important properties of approximations. Therefore, in such models, reducts may not be represented by Boolean functions.

When a measure γ (e.g. Eq. 7.8) representing a part of consistency of a rough set model is given, we can define approximate measure-based reducts as follows: $A \subseteq C$ is an approximate reduct if $\gamma_A \geq (1 - \varepsilon)\gamma_C$ or $\gamma_A \geq \gamma_C - \varepsilon$ for a small $\varepsilon \geq 0$. Several measures used for approximate measure-based reducts have been proposed, e.g. based on the number of discerned object pairs or the information entropy [45, 46, 51]. Comparing with structure-based reducts, approximate measure-based approach can easily control size of reducts, but we cannot expect which parts of the structure of the rough set model deteriorate by reduction.

We show that reducts are (approximately) represented by prime implicants of Boolean functions (or pairs of Boolean functions). To compute all reducts of a particular type, we solve the dualization problem (more precisely, positive DNF (or CNF) dualization) of the corresponding Boolean function. It probably cannot be solved in polynomial time (it can be solved in quasi-polynomial time with respect to the sizes of the input and the output [9]).

To apply attribute reduction of this chapter to real-world data sets, we notice the following three points. Firstly, we need additional measures to select the best reducts for applications, for example, minimizing the size of the reduct or the number of the equivalence classes given by the reduct and so on [1, 13, 27, 42, 48, 50]. Such an optimization problem cannot be generally solved in polynomial time. Therefore, there are heuristic methods computing one or a number of reducts which are near to optimal [1, 27, 42, 48]. It does not mean that the Boolean functions studied in this chapter are useless for applications. They can be incorporated into heuristic methods.

Secondly, when data sets include numerical or continuous attribute values, the approach of this chapter does not work well, because the order of values or the degree of difference between values are not considered (except for criteria in DRSM). There are two approaches to overcome the drawback. One is discretization [7, 15] where the domain of a numerical attribute is partitioned to lower number of values. After discretization, we can apply attribute reduction to the data set without modification. The other is to use a similarity relation [12, 28] instead of the indiscernibility relation or a fuzzy partition [12, 22, 27] instead of the equivalence classes and define extensions of RSM. In that case, we can define structure-based reducts for the extended RSMs in the same way as those of this chapter.

Thirdly, reducts could suffer from overfitting because of rigid definitions of their preserving conditions. One technique to avoid overfitting is dynamic reducts [1], where decision tables with object subsets of a given cardinality are randomly and repeatedly selected, and reducts which appear in more decision tables than a given threshold are chosen as dynamic reducts.

In this chapter, we did not discuss algorithms to compute reducts and numerical experiments, whereas they are found in [5, 6, 8, 11, 13, 17, 19, 34, 41, 50, 51]. The references show how to select a desirable reduct or find an optimal reduct, and how to use the selected reduct for building classifiers. Additionally, they also show experimental results for benchmark or real-world data sets. The references do not include some types of reducts of this chapter, especially most types of reducts in VPRSM, however, from their results we hope that the proposed reducts would be useful in applications.

Proofs of theoretical results of this chapter are not so difficult. Parts of proofs are found in our papers [20, 23, 25, 31].

References

1. Bazan, J.G., Nguyen, H.S., Nguyen, S.H., Synak, P., Wróblewski, J.: Rough set algorithms in classification problem. In: Polkowski, L., Tsumoto, S., Lin, T.Y. (eds.) Rough Set Methods and Applications, pp. 49–88. Physica-Verlag, New York (2000)
2. Ben-David, A.: Monotonicity maintenance in information-theoretic machine learning algorithms. Mach. Learn. **19**, 29–43 (1995)
3. Ben-David, A., Sterling, L., Pao, Y.H.: Learning and classification of monotonic ordinal concepts. Comput. Intell. **5**(1), 45–49 (1989)

4. Beynon, M.: Reducts within the variable precision rough sets model: a further investigation. Eur. J. Oper. Res. **134**(3), 592–605 (2001)

5. Beynon, M.J., Peel, M.J.: Variable precision rough set theory and data discretisation: an application to corporate failure prediction. Omega **29**, 561–576 (2001)

6. Chen, D., Hu, Q., Yang, Y.: Parameterized attribute reduction with Gaussian kernel based fuzzy rough sets. Inf. Sci. **181**, 5169–5179 (2011)

7. Chmielewski, M.R., Grzymala-Busse, J.W.: Global discretization of continuous attributes as preprocessing for machine learning. Int. J. Approx. Reason. **15**, 319–331 (1996)

8. Cornelis, C., Jensen, R., Hurtado, G., Ślęzak, D.: Attribute selection with fuzzy decision reducts. Inf. Sci. **180**, 209–224 (2010)

9. Crama, Y., Hammer, P.L.: Boolean Functions: Theory, Algorithms, and Applications. Cambridge University Press, New York (2011)

10. Dembczyński, K., Greco, S., Kotłowski, W., Słowiński, R.: Quality of rough approximation in multi-criteria classification problems. In: Greco, S., Hata, Y., Hirano, S., Inuiguchi, M., Miyamoto, S., Nguyen, H.S., Słowiński, R. (eds.) 5th International Conference on Rough Sets and Current Trends in Computing, RSCTC 2006. LNCS (LNAI), vol. 4259, pp. 318–327. Springer, Heidelberg (2006)

11. Dimitras, A.I., Slowinski, R., Susmaga, R., Zopounidis, C.: Business failure prediction using rough sets. Eur. J. Oper. Res. **114**, 263–280 (1999)

12. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. Int. J. Gen. Syst. **17**(2–3), 191–209 (1990)

13. Düntsch, I., Gediga, G.: Uncertainty measures of rough set prediction. Artif. Intell. **106**, 109–137 (1998)

14. Eiter, T., Makino, K., Gottlob, G.: Computational aspects of monotone dualization: a brief survey. Discret. Appl. Math. **156**, 2035–2049 (2011)

15. Fayyad, U.M., Irani, K.B.: On the handling of continuous-valued attributes in decision tree generation. Mach. Learn. **8**(1), 87–102 (1992)

16. Greco, S., Matarazzo, B., Slowinski, R.: Rough set theory for multicriteria decision analysis. Eur. J. Oper. Res. **129**(1), 1–47 (2001)

17. Greco, S., Matarazzo, B., Slowinski, R.: Multicriteria classification by dominance-based rough set approach. In: Kloesgen, W., Zytkow, J.M. (eds.) Handbook of Data Mining and Knowledge Discovery. Oxford University Press, New York (2002)

18. Greco, S., Matarazzo, B., Słowiński, R.: Decision rule approach. In: Figueira, J., Greco, S., Ehrgott, M. (eds.) Multiple Criteria Decision Analysis: State of the Surveys, pp. 507–561. Springer, New York (2005)

19. Hassanien, A.E.: Rough set approach for attribute reduction and rule generation: a case of patients with suspected breast cancer. J. Am. Soc. Inf. Sci. Technol. **55**(11), 954–962 (2004)

20. Inuiguchi, M.: Attribute reduction in variable precision rough set model. Int. J. Uncertain. Fuzziness Knowl. Based Syst. **14**(4), 461–479 (2006)

21. Inuiguchi, M.: Structure-based attribute reduction in variable precision rough set models. J. Adv. Comput. Intell. Intell. Inf. **10**(5), 657–665 (2006)

22. Inuiguchi, M., Tanino, T.: New fuzzy rough sets based on certainty qualification. In: Pal, S.K., Polkowski, L., Skowron, A. (eds.) Rough-Neural Computing: Techniques for Computing with Words, pp. 277–296. Springer, Berlin (2004)

23. Inuiguchi, M., Tsurumi, M.: Measures based on upper approximations of rough sets for analysis of attribute importance and interaction. Int. J. Innov. Comput. Inf. Control **2**(1), 1–12 (2006)

24. Inuiguchi, M., Matsumoto, Y.: Refinement of attribute reduction in the classical rough sets toward decision analysis. In: International Workshop on Soft Computing for Knowledge Technology (2008)

25. Inuiguchi, M., Yoshioka, Y.: Several reducts in dominance-based rough set approach. In: Huynh, V.N., Nakamori, Y., Ono, H., Lawry, J., Kreinovich, V., Nguyen, H.T. (eds.) Interval/Probabilistic Uncertainty and Non-classical Logics. Advances in Soft Computing, vol. 46, pp. 163–175. Springer, Berlin (2008)

26. Inuiguchi, M., Yoshioka, Y., Kusunoki, Y.: Variable-precision dominance-based rough set approach and attribute reduction. Int. J. Approx. Reason. **50**(8), 1199–1214 (2009)
27. Jensen, R., Shen, Q.: Semantics-preserving dimensionality reduction: rough and fuzzy-rough-based approaches. IEEE Trans. Knowl. Data Eng. **16**(12), 1457–1471 (2004)
28. Jensen, R., Tuson, A., Shen, Q.: Finding rough and fuzzy-rough set reducts with SAT. Inf. Sci. **255**, 100–120 (2014)
29. Kryszkiewicz, M.: Rough set approach to incomplete information systems. Inf. Sci. **112**, 39–49 (1998)
30. Kryszkiewicz, M.: Comparative study of alternative types of knowledge reduction in inconsistent systems. Int. J. Intell. Syst. **16**, 105–120 (2001)
31. Kusunoki, Y., Inuiguchi, M.: A unified approach to reducts in dominance-based rough set approach. Soft Comput. **14**, 507–515 (2010)
32. Lievens, S., Baets, B.D., Cao-Van, K.: A probabilistic framework for the design of instance-based supervised ranking algorithms in an ordinal setting. Ann. Oper. Res. **163**, 115–142 (2008)
33. Mi, J., Wu, W., Zhang, W.: Approaches to knowledge reduction based on variable precision rough set model. Inf. Sci. **159**, 255–272 (2004)
34. Nguyen, L.G., Nguyen, H.S.: On elimination of redundant attributes in decision tables. In: Proceedings of the Federated Conference on Computer Science and Information Systems, pp. 317–322 (2012)
35. Pawlak, Z.: Rough sets. Int. J. Inf. Comput. Sci. **11**(5), 341–356 (1982)
36. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht (1991)
37. Pawlak, Z., Skowron, A.: Rough sets and boolean reasoning. Inf. Sci. **177**, 41–73 (2007)
38. Pawlak, Z., Skowron, A.: Rough sets: some extensions. Inf. Sci. **177**, 28–40 (2007)
39. Pawlak, Z., Skowron, A.: Rudiments of rough sets. Inf. Sci. **177**, 3–27 (2007)
40. Pawlak, Z., Słowiński, R.: Rough set approach to multi-attribute decision analysis. Eur. J. Oper. Res. **72**, 443–459 (1994)
41. Sawicki, P., Żak, J.: Technical diagnostic of a fleet of vehicles using rough set theory. Eur. J. Oper. Res. **193**, 891–903 (2009)
42. Shen, Q., Jensen, R.: Rough sets, their extensions and applications. Int. J. Autom. Comput. **4**(3), 217–228 (2007)
43. Skowron, A., Rauszer, C.: The discernibility matrix and function in information systems. In: Słowiński, R. (ed.) Intelligent Decision Support: Handbook of Application and Advances of Rough Set Theory, pp. 331–362. Kluwer Academic Publishers, Dordrecht (1992)
44. Skowron, A., Stepaniuk, J.: Tolerance approximation spaces. Fundam. Inform. **27**(2–3), 245–253 (1996)
45. Ślęzak, D.: Various approaches to reasoning with frequency based decision reducts: a survey. In: Polkowski, L., Tsumoto, S., Lin, T.Y. (eds.) Rough Set Methods and Applications, pp. 235–285. Physica-Verlag, New York (2000)
46. Ślęzak, D.: Approximate entropy reducts. Fundam. Inform. **53**, 365–390 (2002)
47. Ślęzak, D., Ziarko, W.: The investigation of the Bayesian rough set model. Int. J. Approx. Reason. **40**, 81–91 (2005)
48. Ślęzak, D., Janusz, A.: Ensembles of bireducts: towards robust classification and simple representation. In: Kim, T.H., Adeli, H., Slezak, D., Sandnes, F.E., Song, X., Chung, K.I., Arnett, K.P. (eds.) Future Generation Information Technology: Third International Conference, FGIT 2011. LNCS, vol. 7105, pp. 64–77. Springer, Berlin (2011)
49. Susmaga, R., Słowiński, R., Greco, S., Matarazzo, B.: Generation of reducts and rules in multi-attribute and multi-criteria classification. Control Cybern. **29**(4), 969–988 (2000)
50. Swiniarski, R.W., Skowron, A.: Rough set methods in feature selection and recognition. Pattern Recognit. Lett. **24**, 833–849 (2003)
51. Wróblewski, J.: Ensembles of classifiers based on approximate reducts. Fundam. Inform. **47**, 351–360 (2001)
52. Yang, X., Yang, J., Wu, C., Yu, D.: Dominance-based rough set approach and knowledge reductions in incomplete ordered information system. Inf. Sci. **178**, 1219–1234 (2008)
53. Ziarko, W.: Variable precision rough set model. J. Comput. Syst. Sci. **46**(1), 39–59 (1993)
54. Ziarko, W.: Probabilistic approach to rough sets. Int. J. Approx. Reason. **49**, 272–284 (2008)