Probabilistic

24. Probabilistic Rough Sets

Yiyu Yao, Salvatore Greco, Roman Słowiński

As quantitative generalizations of Pawlak rough sets, probabilistic rough sets consider degrees of overlap between equivalence classes and the set. An equivalence class is put into the lower approximation if the conditional probability of the set, given the equivalence class, is equal to or above one threshold; an equivalence class is put into the upper approximation if the conditional probability is above another threshold hold. We review a basic model of probabilistic rough sets (i. e., decision-theoretic rough set model) and variations. We present the main results of probabilistic rough sets by focusing on three issues: (a) interpretation and calculation of the required thresholds, (b) estimation of the required conditional probabilities, and (c) interpretation and applications of probabilistic rough set approximations.

24.1 Motivation for Studying Probabilistic Rough Sets

Rough set theory [24.1, 2] provides a simple and elegant method for analyzing data represented in a tabular form called an information table. The rows of the table represent a finite set of objects, the columns represent a finite set of attributes, and each cell represents the value of an object on the corresponding attribute. With a limited number of attributes, we may only be able to describe some subsets of objects precisely [24.3, 4]. Those subsets that can be precisely described are called definable sets, and all other subsets are called undefinable sets. A fundamental notion of rough set theory is the approximation of a subset of objects by a pair of definable sets from below and above, or equivalently, by three pairwise disjoint positive, negative, and boundary regions [24.4].

Pawlak rough set approximations are characterized by a zero tolerance of errors. That is, an object in the lower approximation certainly belongs to set and an object in the complement of the upper approximation certainly does not belong to the set. This has motivated the introduction of many different generalizations of rough sets. By introducing certain levels of errors, probabilistic rough sets [24.5, 6] are quantitative generalizations of the qualitative Pawlak rough sets. Although several specific models of probabilistic rough sets had been considered by some authors [24.7–10], a more general model, called decision-theoretic rough set (DTRS) model, was first proposed by *Yao* et al. [24.11, 12] based on the well-established Bayesian decision theory. Other probabilistic models include variable precision rough sets [24.13, 14], Bayesian rough sets [24.15–18], parameterized rough sets [24.19, 20], game-theoretic rough sets [24.21, 22], variable-consistency-indiscernibility-based and dominance-based rough sets [24.23, 24], stochastic dominance-based rough sets [24.25], naive Bayesian rough sets [24.26], information-theoretic rough sets [24.27], confirmation-theoretic rough sets [24.28], and many different types of probabilistic rough set approximations [24.29, 30].

In this chapter, we present a basic model of probabilistic rough sets and a brief review of other probabilistic rough set models. We examine in particular three fundamental issues, namely, the interpretation and computation of the pair of thresholds, the estimation of probability, and an application of three regions. We also show how a probabilistic approach can be applied when information related to some order representing the extent to which some property related to considered attributes has to be taken into account. This situation is handled by the well-known rough set extension called dominance-based rough set approach [24.31–34]. A full understanding of these issues will greatly increase the chance of success when applying probabilistic rough sets in real-world applications.

24.2 Pawlak Rough Sets

We present a semantically meaningful definition of rough set approximations and a simple method for constructing rough set approximations.

24.2.1 Rough Set Approximations

In rough set theory, a finite set of objects is described by using a finite set of attributes in a tabular form, called an information table [24.2]. Formally, an information table can be expressed as

$$S = (U, AT, \{V_a \mid a \in AT\}, \{I_a \mid a \in AT\}) ,$$

where

U is a finite nonempty set of objects called universe ,
AT is a finite nonempty set of attributes ,
V_a is a nonempty set of values for $a \in AT$,
$I_a : U \to V_a$ is an information function .

The information table provides all available information about the set of objects, based on which we can perform tasks of analysis and inference.

In an information table, one can introduce a description language, as suggested by *Marek* and *Pawlak* [24.3], to formally describe objects. We consider a language DL that is recursively defined as follows

$$\begin{aligned} &(1)\ \ (a = v) \in DL,\ \text{where } a \in AT, v \in V_a ,\\ &(2)\ \ \text{if } p, q \in DL,\ \text{then } \neg p, p \wedge q, p \vee q \in DL . \end{aligned}$$

Formulas defined by (1) are called atomic formulas. The satisfiability of formula p by an object x, written

$x \models p$, is defined as follows

(i) $x \models (a = v)$, iff $I_a(x) = v$,
(ii) $x \models \neg p$, iff $\neg(x \models p)$,
(iii) $x \models p \wedge q$, iff $x \models p$ and $x \models q$,
(iv) $x \models p \vee q$, iff $x \models p$ or $x \models q$.

If p is a formula, the set $m(p) \subseteq U$ defined by

$$m(p) = \{x \in U \mid x \models p\} \tag{24.1}$$

is called the meaning set of p. That is, the meaning set $m(p)$ consists of all those objects that satisfy the formula p.

With the introduction of a description language, we can formally describe an important characteristics of an information table, namely, some subsets of objects are definable or describable while others are not. A subset of objects $X \subseteq U$ is called a definable set [24.3, 4] if there exists a formula p such that

$$X = m(p) , \tag{24.2}$$

otherwise, X is called an undefinable set. The formula p is called a description of X. Let $\mathrm{DEF}(U) \subseteq 2^U$ denote the family of all definable sets, where 2^U is the power set of U. By definition, $\mathrm{DEF}(U)$ contains the empty set $\emptyset$, the entire universe U and is closed under set complement, intersection, and union. In other words, $\mathrm{DEF}(U)$ is a sub-Boolean algebra of the power set 2^U.

For any subset of objects $X \subseteq U$, may be either definable or undefinable, we define the following pair of lower and upper approximations

$$\begin{aligned}
\underline{apr}(X) &= \text{the largest definable set contained by } X \\
&= \bigcup\{G \in \mathrm{DEF}(U) | G \subseteq X\} , \\
\overline{apr}(X) &= \text{the smallest definable set containing } X \\
&= \bigcap\{G \in \mathrm{DEF}(U) | X \subseteq G\} .
\end{aligned} \tag{24.3}$$

By definition, it follows that $\underline{apr}(X) \subseteq X \subseteq \overline{apr}(X)$ for any $X \subseteq U$, and $\underline{apr}(X) = X = \overline{apr}(X)$ if and only if $X \in \mathrm{DEF}(U)$. The definition is semantically meaningful in the sense that it clearly explains the motivation for introducing rough set approximations and provides an interpretation of the approximations. However, one cannot use this definition to construct rough set approximations easily.

24.2.2 Construction of Rough Set Approximations

A simple method for constructing rough set approximations is through an equivalence relation. For an attribute $a \in AT$, the information function I_a maps an object in U to a value of V_a, that is, $I_a(x) \in V_a$. For an attribute $a \in AT$, we can define an equivalence relation E_a as follows: for $x, y \in U$

$$xE_a y \Longleftrightarrow I_a(x) = I_a(y) . \tag{24.4}$$

The equivalence class containing x is denoted by $[x]_a$. Similarly, for a subset of attributes $A \subseteq AT$, we define an equivalence relation E_A

$$xE_A y \Longleftrightarrow \forall a \in A(I_a(x) = I_a(y)) . \tag{24.5}$$

The equivalence class containing x is denoted by $[x]_A$. By definition, it follows that, for $a \in AT$ and $A \subseteq AT$,

$$\begin{aligned}
&E_{\{a\}} = E_a , \qquad [x]_{\{a\}} = [x]_a , \\
&E_A = \bigcap_{a \in A} E_a , \quad [x]_A = \bigcap_{a \in A} [x]_a .
\end{aligned} \tag{24.6}$$

That is, we can construct the equivalence relation induced by a subset of attributes A by using equivalence relations induced by individual attributes in A.

Consider the equivalence relation $E_A \subseteq U \times U$ induced by a subset of attributes $A \subseteq AT$. The equivalence relation E_A induces a partition U/E_A of U, i. e., a family of nonempty and pairwise disjoint subsets whose union is the universe. For an object $x \in U$, its equivalence class is given by

$$[x]_A = \{y \in U \mid xE_A y\} . \tag{24.7}$$

By taking the union of a family of equivalence classes, one can construct an atomic sub-Boolean $B(U/E_A)$ of 2^U with U/E_A as the set of atoms

$$B(U/E_A) = \{\bigcup F \mid F \subseteq U/E_A\} . \tag{24.8}$$

That is, $B(U/E_A)$ contains the empty set $\emptyset$, the whole set U, and is closed with respect to set complement, intersection, and union. The three notions of equivalence relation E, the partition U/E_A, and atomic Boolean algebra $B(U/E_A)$ uniquely determine each other. We can therefore use E_A, U/E_A, and $B(U/E_A)$ interchangeably.

The pair $apr = (U, E_A)$, equivalently, the pair $apr = (U, U/E_A)$ or the pair $apr = (U, B(U/E_A))$, is called an approximation space. Although three different representations are equivalent, each of them provides a different hint when we generalize rough sets. The pair $apr = (U, E_A)$ is useful for generalizing rough sets using a nonequivalence relation [24.35]. The partition U/E_A may be viewed as a granulation of the universe U and the pair $apr = (U, U/E_A)$ relates rough sets and granular computing [24.36]. The pair $apr = (U, B(U/E_A))$ leads to a subsystem-based formulation and generalizations [24.37].

For a subset of attributes $A \subseteq AT$, if we restrict the formulas of DL by using only attributes in A, we obtain a sublanguage $DL(A) \subseteq DL$. It can be proved that the family of all definable sets $\mathrm{DEF}_A(U)$ defined by $DL(A)$ is exactly the sub-Boolean algebra $B(U/E_A)$. With respect to a subset of attributes $A \subseteq AT$, each object x is described by a logic formula

$$\bigwedge_{a \in A} a = I_a(x) , \tag{24.9}$$

where $I_a(x) \in V_a$ and the atomic formula $a = I_a(x)$ indicate that the value of an object on attribute a is $I_a(x)$. The equivalence class containing x, namely, $[x]_{E_A}$, is the set of those objects that satisfy the formula $\wedge_{a \in A} a = I_a(x)$. The formula can be viewed as a description of objects that are equivalent to x with respect to A, including x itself.

Based on the equivalence of $\mathrm{DEF}_A(U)$ and $B(U/E_A)$, we can equivalently define rough set approximations by using the equivalence classes $[x]_A$. For simplicity, we also simply write $[x]$ when no confusion arises.

For a subset of objects $X \subseteq U$, the pair of lower and upper approximations can be equivalently defined by

$$\begin{aligned} \underline{apr}(X) &= \{x \in U \mid [x] \subseteq X\} , \\ \overline{apr}(X) &= \{x \in U \mid [x] \cap X \neq \emptyset\} . \end{aligned} \tag{24.10}$$

Construction of rough set approximation by this definition is much easier. Alternatively, one can also define three pairwise disjoint positive, negative, and boundary regions [24.38]

$$\begin{aligned} \mathrm{POS}(X) &= \{x \in U \mid [x] \subseteq X\} , \\ \mathrm{NEG}(X) &= \{x \in U \mid [x] \cap X = \emptyset\} , \\ \mathrm{BND}(X) &= \{x \in U \mid [x] \nsubseteq X \wedge [x] \cap X \neq \emptyset\} . \end{aligned} \tag{24.11}$$

The pair of approximations and three regions determines each other as follows

$$\begin{aligned} \mathrm{POS}(X) &= \underline{apr}(X) , \\ \mathrm{NEG}(X) &= (\overline{apr}(X))^c , \\ \mathrm{BND}(X) &= \overline{apr}(X) - \underline{apr}(X) , \end{aligned} \tag{24.12}$$

and

$$\begin{aligned} \underline{apr}(X) &= \mathrm{POS}(X) , \\ \overline{apr}(X) &= \mathrm{POS}(X) \cup \mathrm{BND}(X) , \end{aligned} \tag{24.13}$$

where $(\cdot)^c$ denotes the complement of a set. Each representation provides a distinctive interpretation of rough set approximations. We will use the three-region approximation in the rest of this chapter, due to its close connections to three-way decisions.

24.3 A Basic Model of Probabilistic Rough Sets

Decision-theoretic rough set (DTRS) model proposed by *Yao* et al. [24.11, 12] gives rises to a general form of probabilistic rough set approximations by using a pair of thresholds on conditional probabilities. The results enable us to formulate a basic model of probabilistic rough sets. However, we introduce the model in a way that is different from DTRS. We first interpret Pawlak rough sets in terms of probability and the two extreme value of probability (i. e., 1 and 0) and then generalize 1 and 0 into a pair of thresholds (α, β) with $0 \leq \beta < \alpha \leq 1$.

The Pawlak rough sets consider only qualitative relationship between an equivalence class and a set, namely, an equivalence is a subset of the set or has a nonempty intersection with the set. This qualitative nature becomes clearer with a probabilistic interpretation [24.6]. Suppose $Pr(X|[x])$ denotes the conditional probability that an object is in X given that the object is in $[x]$. The conditions for defining rough set three regions can be equivalently expressed as

$$\begin{aligned} &[x] \subseteq X \Longleftrightarrow Pr(X|[x]) \geq 1 ; \\ &[x] \cap X = \emptyset \Longleftrightarrow Pr(X|[x]) \leq 0 ; \\ &[x] \nsubseteq X \wedge [x] \cap X \neq \emptyset \Longleftrightarrow 0 < Pr(X|[x]) < 1 . \end{aligned} \tag{24.14}$$

Although a probability can never be greater than 1 or less than 0, we purposely use the conditions ≥ 1 and

≤ 0 whose intended meaning will become clearer later. By those conditions, Pawlak three regions can be equivalently expressed as

$$\begin{aligned}
\mathrm{POS}(X) &= \{x \in U \mid Pr(X|[x]) \geq 1\} ,\\
\mathrm{NEG}(X) &= \{x \in U \mid Pr(X|[x]) \leq 0\} ,\\
\mathrm{BND}(X) &= \{x \in U \mid 0 < Pr(X|[x]) < 1\} . \qquad (24.15)
\end{aligned}$$

They show that Pawlak rough sets only use the two extreme values, i. e., 1 and 0, of probability.

It is natural to generalize Pawlak rough sets by replacing 1 and 0 with some other values in the unit interval $[0, 1]$. Given a pair of thresholds α, β with $0 \leq \beta < \alpha \leq 1$, the main results of probabilistic rough sets are the (α, β)-probabilistic regions defined by

$$\begin{aligned}
\mathrm{POS}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|[x]) \geq \alpha\} ,\\
\mathrm{NEG}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|[x]) \leq \beta\} ,\\
\mathrm{BND}_{(\alpha,\beta)}(X) &= \{x \in U \mid \beta < Pr(X|[x]) < \alpha\} . \qquad (24.16)
\end{aligned}$$

The Pawlak rough set model is a special case in which $\alpha = 1$ and $\beta = 0$. In the case when $0 < \beta = \alpha < 1$, the three regions are given by

$$\begin{aligned}
\mathrm{POS}_{(\alpha,\alpha)}(X) &= \{x \in U \mid Pr(X|[x]) > \alpha\} ,\\
\mathrm{NEG}_{(\alpha,\alpha)}(X) &= \{x \in U \mid Pr(X|[x]) < \alpha\} ,\\
\mathrm{BND}_{(\alpha,\alpha)}(X) &= \{x \in U \mid Pr(X|[x]) = \alpha\} . \qquad (24.17)
\end{aligned}$$

It may be commented that this special case is perhaps more of mathematical interest, rather than practical applications. We use this particular definition in order to establish connection to existing studies. As will be shown in subsequent discussions, when $Pr(X|[x]) = \alpha = \beta$, the costs of assigning objects in $[x]$ to the positive, boundary, and negative regions, respectively, are the same. In fact, one may simply define two regions by assigning objects in the boundary region into either the positive or boundary region.

The main results of the basic model of probabilistic rough sets were first proposed by *Yao* et al. [24.11, 12] in a DTRS model, based on Bayesian decision theory. The DTRS model covers all specific models introduced before it. The interpretation of Pawlak rough sets in terms of conditional probability, i. e., the model characterized by $\alpha = 1$ and $\beta = 0$, was first given by *Wong* and *Ziarko* [24.10]. A 0.5-model, characterized by $\alpha = \beta = 0.5$, was introduced by *Wong* and *Ziarko* [24.8] and *Pawlak* et al. [24.7], in which the positive region is defined by probability greater than 0.5, the negative by probability less than 0.5, and the boundary by probability equal to 0.5. A model characterized by $\alpha > 0.5$ and $\beta = 0.5$ was suggested by *Wong* and *Ziarko* [24.9]. Most recent developments on decision-theoretic rough sets can be found in a book edited by *Li* et al. [24.39] and papers [24.21, 40–50] in a journal special issue edited by *Yao* et al. [24.51].

24.4 Variants of Probabilistic Rough Sets

Since the introduction of decision-theoretic rough set model, several new models have been proposed and investigated. They offer related but different directions in generalizing Pawlak rough sets by incorporating probabilistic information.

24.4.1 Variable Precision Rough Sets

The first version of variable precision rough sets was introduced by *Ziarko* [24.14], in which the standard set inclusion $[x] \subseteq X$ is generalized into a graded set inclusion $s([x], X)$ called a measure of the relative degree of misclassification of $[x]$ with respect to X. A particular measure suggested by Ziarko is given by

$$s([x], X) = 1 - \frac{|[x] \cap X|}{|[x]|} , \qquad (24.18)$$

where $|\cdot|$ denotes the cardinality of a set. By introducing a threshold $0 \leq z < 0.5$, one can define three regions as follows

$$\begin{aligned}
\mathrm{VPOS}_z(X) &= \{x \in U \mid s([x], X) \leq z\} ,\\
\mathrm{VNEG}_z(X) &= \{x \in U \mid s([x], X) \geq 1 - z\} ,\\
\mathrm{VBND}_z(X) &= \{x \in U \mid z < s([x], X) < 1 - z\} . \qquad (24.19)
\end{aligned}$$

A more generalized version using a pair of thresholds was late introduced by *Katzberg* and *Ziarko* [24.13] as follows: for $0 \leq l < u \leq 1$,

$$\begin{aligned}
\mathrm{VPOS}_{(l,u)}(X) &= \{x \in U \mid s([x], X) \leq l\} ,\\
\mathrm{VNEG}_{(l,u)}(X) &= \{x \in U \mid s([x], X) \geq u\} ,\\
\mathrm{VBND}_{(l,u)}(X) &= \{x \in U \mid l < s([x], X) < u\} . \qquad (24.20)
\end{aligned}$$

The one-threshold model may be considered as a special case of the two-threshold model with $l = z$ and $u = 1 - z$.

One may interpret the ratio in (24.18) as an estimation of the conditional probability $Pr(X|[x])$, namely

$$s([x], X) = 1 - \frac{|[x] \cap X|}{|[x]|} = 1 - Pr(X|[x]) \ . \quad (24.21)$$

By setting $\alpha = 1 - l$ and $\beta = 1 - u$, we immediately have

$$\begin{aligned} \mathrm{POS}_{(1-l,1-u)}(X) &= \{x \in U \mid Pr(X|[x]) \geq 1 - l\} \\ &= \{x \in U \mid s([x], X) \leq l\} \\ &= \mathrm{VPOS}_{(l,u)}(X) \ , \\ \mathrm{NEG}_{(1-l,1-u)}(X) &= \{x \in U \mid Pr(X|[x]) \geq 1 - u\} \\ &= \{x \in U \mid s([x], X) \geq u\} \\ &= \mathrm{VNEG}_{(l,u)}(X) \ , \\ \mathrm{BND}_{(1-l,1-u)}(X) &= \{x \in U \mid 1 - l < Pr(X|[x]) < 1 - u\} \\ &= \{x \in U \mid l < s([x], X) < u\} \\ &= \mathrm{VBND}_{(l,u)}(X) \ . \end{aligned} \quad (24.22)$$

It follows that, when the particular set-inclusion measure defined by (24.18) is used, the variable precision rough sets are coincident with the decision-theoretic rough sets.

Variable precision rough sets provide an alternative direction in generalizing Pawlak rough sets by considering a graded set-inclusion relation, which is not necessarily restricted to a probabilistic interpretation. If we use other set-inclusion measures, we will obtain other types of quantitative rough sets [24.38, 52]. Unfortunately, subsequent developments lose this crucial feature in an attempt to unify variable precision rough sets into probabilistic rough sets [24.53].

24.4.2 Parameterized Rough Sets

Parameterized rough sets, proposed by *Greco* et al. [24.19, 20], generalize probabilistic rough sets by introducing a Bayesian confirmation measure and a pair of thresholds on the confirmation measure, in addition to a pair of thresholds on conditional probability. According to *Fitelson* [24.54], *measures of confirmation* quantify the degree to which a piece of evidence E provides *evidence for or against* or *support for or against* a hypothesis H.

A measure of confirmation of a piece of evidence E with respect to a hypothesis H is denoted by $c(E, H)$. A confirmation measure $c(E, H)$ is required to satisfy the following minimal property:

$$c(E, H) = \begin{cases} > 0 & \text{if } Pr(H|E) > Pr(H) \\ = 0 & \text{if } Pr(H|E) = Pr(H) \\ < 0 & \text{if } Pr(H|E) < Pr(H) . \end{cases} \quad (i)$$

Two well-known Bayesian confirmation measures are [24.55]

$$\begin{aligned} c_d([x], X) &= Pr(X|[x]) - Pr(X) \ , \\ c_r([x], X) &= \frac{Pr(X|[x])}{Pr(X)} \ . \end{aligned} \quad (24.23)$$

These measures have a probabilistic interpretation. The parameterized rough sets can be therefore viewed as a different formulation of probabilistic rough sets.

A first discussion about relationships between confirmation measures and rough sets were proposed by *Greco* et al. [24.56]. Other contributions related to the properties of confirmation measures with special attention to application to rough sets are given in [24.57].

Given a pair of thresholds (s, t) with $t < s$, three (α, β, s, t)-parameterized regions are defined in [24.19, 20]

$$\begin{aligned} \mathrm{PPOS}_{(\alpha,\beta,s,t)}(X) &= \{x \in U \mid Pr(X|[x]) \geq \alpha \\ &\quad \wedge c([x], X) \geq s\} \ , \\ \mathrm{PNEG}_{(\alpha,\beta,s,t)}(X) &= \{x \in U \mid Pr(X|[x]) \leq \beta \\ &\quad \wedge c([x], X) \leq t\} \ , \\ \mathrm{PBND}_{(\alpha,\beta,s,t)}(X) &= \{x \in U \mid (Pr(X|[x]) > \beta \\ &\quad \vee c([x], X) > t) \\ &\quad \wedge (Pr(X|[x]) < \alpha \\ &\quad \vee c([x], X) < s)\} \ . \end{aligned} \quad (24.24)$$

There exist many Bayesian confirmation measures, which makes the model of parameterized rough sets more flexible. On the other hand, due to lack of a general agreement on a Bayesian confirmation measure, choosing an appropriate confirmation measure for a particular application may not be an easy task.

24.4.3 Confirmation-Theoretic Rough Sets

Although many Bayesian confirmation measures are related to the conditional probability $Pr(X|[x])$, *Zhou* and *Yao* [24.26] argued that the conditional probability $Pr(X|[x])$ and a Bayesian confirmation measure have very different semantics and should be used for different purposes. For example, the conditional probability $Pr(X|[x])$ gives us an absolute degree of confidence in classifying objects from $[x]$ as belonging to X. On the other hand, a Bayesian measure, for example, c_d or c_r, normally reflects a change of confidence in X before and after knowing $[x]$. Thus, a Bayesian confirmation measures is useful to weigh the strength of evidence $[x]$ with respect to the hypothesis X. A mixture of conditional probability and confirmation measure in the parameterized rough sets may cause a semantic difficulty in interpreting the three regions.

To resolve this difficulty, *Zhou* and *Yao* [24.28] suggested a separation of the parameterized model into two models. One is the conventional probabilistic model and the other is a confirmation-theoretic model. For a Bayesian confirmation measure $c([x], X)$ and a pair of thresholds (s, t) with $t < s$, three confirmation regions are defined by

$$\begin{aligned}
\mathrm{CPOS}_{(s,t)}(X) &= \{[x] \in U/R \mid c([x], X) \geq s\}\,, \\
\mathrm{CNEG}_{(s,t)}(X) &= \{[x] \in U/R \mid c([x], X) \leq t\}\,, \\
\mathrm{CBND}_{(s,t)}(X) &= \{[x] \in U/R \mid t < c([x], X) < s\}\,.
\end{aligned} \tag{24.25}$$

For the case with $s = t$, we define

$$\begin{aligned}
\mathrm{CPOS}_{(s,s)}(X) &= \{[x] \in U/R \mid c([x], X) > s\}\,, \\
\mathrm{CNEG}_{(s,s)}(X) &= \{[x] \in U/R \mid c([x], X) < s\}\,, \\
\mathrm{CBND}_{(s,s)}(X) &= \{[x] \in U/R \mid c([x], X) = s\}\,.
\end{aligned} \tag{24.26}$$

In the definition, each equivalence class may be viewed as a piece of evidence. Thus, the partition U/E, instead of the universe, is divided into three regions. An equivalence class in the positive region supports X to a degree at least s, an equivalence class in the negative region supports X to a degree at most t and may be viewed as against X, and an equivalence class in the boundary region is interpreted as neutral toward X.

24.4.4 Bayesian Rough Sets

Bayesian rough sets were proposed by *Ślęzak* and *Ziarko* [24.15, 16] as a probabilistic model in which the required pair of thresholds is interpreted using the a priori probability $Pr(X)$. They introduced Bayesian rough sets and variable precision Bayesian rough sets.

For the Bayesian rough sets, the three regions are defined by

$$\begin{aligned}
\mathrm{BPOS}(X) &= \{x \in U \mid Pr(X|[x]) > Pr(X)\}\,, \\
\mathrm{BNEG}(X) &= \{x \in U \mid Pr(X|[x]) < Pr(X)\}\,, \\
\mathrm{BBND}(X) &= \{x \in U \mid Pr(X|[x]) = Pr(X)\}\,.
\end{aligned} \tag{24.27}$$

Bayesian rough sets can be viewed as a special case of the decision-theoretic rough sets when $\alpha = \beta = Pr(X)$. Semantically, they are very different, however. In contrast to decision-theoretic rough sets, for a set with a higher a priori probability $Pr(X)$, many equivalence classes may not be put into the positive region in the Bayesian rough set model, as the condition $Pr(X|[x]) > Pr(X)$ may not hold. For example, the positive region of the entire universe is always empty, namely, $\mathrm{BPOS}(U) = \emptyset$. This leads to a difficulty in interpreting the positive region as a lower approximation of a set.

The difficulty with Bayesian rough sets stems from the fact that they are in fact a special model of confirmation-theoretic rough sets, which is suitable for classifying pieces of evidence (i. e., equivalence classes), but is inappropriate for approximating a set. Recall that one Bayesian confirmation measure is given by $c_d([x], X) = Pr(X|[x]) - Pr(X)$. Therefore, Bayesian rough sets can be expressed as confirmation-theoretic rough sets as follows,

$$\begin{aligned}
\mathrm{BPOS}(X) &= \{x \in U \mid Pr(X|[x]) > Pr(X)\}\,, \\
&= \{x \in U \mid c_d([x], X) > 0\}\,, \\
&= \bigcup \mathrm{CPOS}_{(0,0)}(X)\,, \\
\mathrm{BNEG}(X) &= \{x \in U \mid Pr(X|[x]) < Pr(X)\}\,, \\
&= \{x \in U \mid c_d([x], X) < 0\}\,, \\
&= \bigcup \mathrm{CNEG}_{(0,0)}(X)\,, \\
\mathrm{BBND}(X) &= \{x \in U \mid Pr(X|[x]) = Pr(X)\} \\
&= \{x \in U \mid c_d([x], X) = 0\}\,, \\
&= \bigcup \mathrm{CBND}_{(0,0)}(X)\,.
\end{aligned} \tag{24.28}$$

That is, the Bayesian rough sets are a model of confirmation-theoretic rough sets characterized by the Bayesian confirmation measure c_d with a pair of thresholds $s = t = 0$. *Ślęzak* and *Ziarko* [24.17] showed that

Bayesian rough sets can also be interpreted by using other Bayesian confirmation measures.

The three regions of the variable precision Bayesian rough sets are defined as follows [24.16]: for $\epsilon \in [0, 1)$

$$\begin{aligned} \mathrm{VBPOS}_{\epsilon}(X) &= \{x \in U \mid Pr(X|[x]) \\ &\qquad \geq 1 - \epsilon(1 - Pr(X))\} , \\ \mathrm{VBNEG}_{\epsilon}(X) &= \{x \in U \mid Pr(X|[x]) \leq \epsilon Pr(X)\} , \\ \mathrm{VBBND}_{\epsilon}(X) &= \{x \in U \mid \epsilon Pr(X) < Pr(X|[x]) \\ &\qquad < 1 - \epsilon(1 - Pr(X))\} . \end{aligned} \tag{24.29}$$

Consider the Bayesian confirmation measure

$$c_r([x], X) = Pr(X|[x])/Pr(X) .$$

For the condition of the positive region, when $Pr(X^c) \neq 0$ we have

$$Pr(X|[x]) \geq 1 - \epsilon(1 - Pr(X)) \Longleftrightarrow c_r([x], X^c) \leq \epsilon . \tag{24.30}$$

Similarly, for the condition defining the negative region, when $Pr(X) \neq 0$ we have

$$Pr(X|[x]) \leq \epsilon Pr(X) \Longleftrightarrow c_r([x], X) \leq \epsilon . \tag{24.31}$$

That is, $[x]$ is put into the positive region if it confirms X^c to a degree less than or equal to ϵ and is put into the negative region if it confirms X to a degree less than or equal to ϵ. In this way, we get a confirmation-theoretic interpretation of variable precision Bayesian rough sets.

Unlike the confirmation-theoretic model defined by (24.25), the positive region of variable precision Bayesian rough sets is defined based on the confirmation of the complement of X and negative region is defined based on the confirmation of X. This definition is a bit awkward to interpret. Generally speaking, it may be more natural to define the positive region by those equivalence classes that confirm X to at least a certain degree. This suggests that one can redefine variable precision Bayesian rough sets by using the framework of confirmation-theoretic rough sets. Moreover, one can use a pair of thresholds instead of one threshold.

24.5 Three Fundamental Issues of Probabilistic Rough Sets

For practical applications of probabilistic rough sets, one must consider at least the following three fundamental issues [24.58, 59]:

- Interpretation and determination of the required pair of thresholds,
- Estimation of the required conditional probabilities, and
- Interpretation and applications of three probabilistic regions.

For each of the three issues, this section reviews one example of the possible methods.

24.5.1 Decision-Theoretic Rough Set Model: Determining the Thresholds

A decision-theoretic model formulates the construction of rough set approximations as a Bayesian decision problem with a set of two states and a set of three actions [24.11, 12]. The set of states is given by $\Omega = \{X, X^c\}$ indicating that an element is in X and not in X, respectively. For simplicity, we use the same symbol to denote both a subset X and the corresponding state. Corresponding to the three regions, the set of actions is given by $\mathcal{A} = \{a_P, a_B, a_N\}$, denoting the actions in classifying an object x, namely, deciding $x \in \mathrm{POS}(X)$, deciding $x \in \mathrm{BND}(X)$, and deciding $x \in \mathrm{NEG}(X)$, respectively. The losses regarding the actions for different states are given by the 3×2 matrix

	X (P)	X^c (N)
a_P	λ_{PP}	λ_{PN}
a_B	λ_{BP}	λ_{BN}
a_N	λ_{NP}	λ_{NN}

In the matrix, λ_{PP}, λ_{BP}, and λ_{NP} denote the losses incurred for taking actions a_P, a_B, and a_N, respectively, when an object belongs to X, and λ_{PN}, λ_{BN} and λ_{NN} denote the losses incurred for taking the same actions when the object does not belong to X.

The expected losses associated with taking different actions for objects in $[x]$ can be expressed as

$$\begin{aligned} R(a_P|[x]) &= \lambda_{PP} Pr(X|[x]) + \lambda_{PN} Pr(X^c|[x]) , \\ R(a_B|[x]) &= \lambda_{BP} Pr(X|[x]) + \lambda_{BN} Pr(X^c|[x]) , \\ R(a_N|[x]) &= \lambda_{NP} Pr(X|[x]) + \lambda_{NN} Pr(X^c|[x]) . \end{aligned} \tag{24.32}$$

The Bayesian decision procedure suggests the following minimum-risk decision rules

(P) If $R(a_P|[x]) \leq R(a_B|[x])$
and $R(a_P|[x]) \leq R(a_N|[x])$, decide $x \in \mathrm{POS}(X)$;
(B) If $R(a_B|[x]) \leq R(a_P|[x])$
and $R(a_B|[x]) \leq R(a_N|[x])$, decide $x \in \mathrm{BND}(X)$;
(N) If $R(a_N|[x]) \leq R(a_P|[x])$
and $R(a_N|[x]) \leq R(a_B|[x])$, decide $x \in \mathrm{NEG}(X)$.

In order to make sure that the three regions are mutually disjoint, tie-breaking criteria should be added when two or three actions have the same risk. We use the following ordering for breaking a tie: a_P, a_N, a_B.

Consider a special class of loss functions with

$$\text{(c0)} \quad \lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}, \quad \lambda_{NN} \leq \lambda_{BN} < \lambda_{PN} . \tag{24.33}$$

That is, the loss of classifying an object x belonging to X into the positive region $\mathrm{POS}(X)$ is less than or equal to the loss of classifying x into the boundary region $\mathrm{BND}(X)$, and both of these losses are strictly less than the loss of classifying x into the negative region $\mathrm{NEG}(X)$. The reverse order of losses is used for classifying an object not in X. With the condition (c0) and the equation $Pr(X|[x]) + Pr(X^c|[x]) = 1$, we can express the decision rules (P)–(N) in the following simplified form (for a detailed derivation, see references [24.58])

(P) If $Pr(X|[x]) \geq \alpha$
and $Pr(X|[x]) \geq \gamma$, decide $x \in \mathrm{POS}(X)$;
(B) If $Pr(X|[x]) \leq \alpha$
and $Pr(X|[x]) \geq \beta$, decide $x \in \mathrm{BND}(X)$;
(N) If $Pr(X|[x]) \leq \beta$
and $Pr(X|[x]) \leq \gamma$, decide $x \in \mathrm{NEG}(X)$,

where

$$\begin{aligned} \alpha &= \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})} , \\ \beta &= \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})} , \\ \gamma &= \frac{(\lambda_{PN} - \lambda_{NN})}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})} . \end{aligned} \tag{24.34}$$

Each rule is defined by two out of the three parameters. By setting $\alpha > \beta$, namely

$$\begin{aligned} &\frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})} \\ &> \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})} , \end{aligned} \tag{24.35}$$

we obtain the following condition on the loss function [24.58]

$$\text{(c1)} \quad \frac{\lambda_{NP} - \lambda_{BP}}{\lambda_{BN} - \lambda_{NN}} > \frac{\lambda_{BP} - \lambda_{PP}}{\lambda_{PN} - \lambda_{BN}} . \tag{24.36}$$

The condition (c1) implies that $1 \geq \alpha > \gamma > \beta \geq 0$. In this case, after tie-breaking, we have the simplified rules [24.58]

(P) If $Pr(X|[x]) \geq \alpha$, decide $x \in \mathrm{POS}(X)$;
(B) If $\beta < Pr(X|[x]) < \alpha$, decide $x \in \mathrm{BND}(X)$;
(N) If $Pr(X|[x]) \leq \beta$, decide $x \in \mathrm{NEG}(X)$.

The parameter γ is no longer needed. Each object can be put into one and only one region by using rules (P), (B), and (N). The (α, β)-probabilistic positive, negative and boundary regions are given, respectively, by

$$\begin{aligned} \mathrm{POS}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|[x]) \geq \alpha\} , \\ \mathrm{BND}_{(\alpha,\beta)}(X) &= \{x \in U \mid \beta < Pr(X|[x]) < \alpha\} , \\ \mathrm{NEG}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|[x]) \leq \beta\} . \end{aligned} \tag{24.37}$$

The formulation provides a solid theoretical basis and a practical interpretation of the probabilistic rough sets. The threshold parameters are systematically calculated from a loss function.

In the development of decision-theoretic rough sets, we assume that a loss function is given by experts in a particular application. There are studies on other types of loss functions and their acquisition [24.60]. Several other proposals have also been made regarding the interpretation and computation of the thresholds, including game-theoretic rough sets [24.21, 22], information-theoretic rough sets [24.27], and an optimization-based framework [24.43, 61, 62].

24.5.2 Naive Bayesian Rough Set Model: Estimating the Conditional Probability

Naive Bayesian rough set model was proposed by *Yao* and *Zhou* [24.59] as a practical method for estimating the conditional probability. First, we perform the logit transformation of the conditional probability

$$\begin{aligned}\operatorname{logit}(Pr(X|[x])) &= \log \frac{Pr(X|[x])}{1-Pr(X|[x])} \\ &= \log \frac{Pr(X|[x])}{Pr(X^c|[x])} ,\end{aligned} \tag{24.38}$$

which is a monotonically increasing transformation of $Pr(X|[x])$. Then, we apply the Bayes' theorem

$$Pr(X|[x]) = \frac{Pr([x]|X)Pr(X)}{Pr([x])} , \tag{24.39}$$

to infer the a posteriori probability $Pr(X|[x])$ from the likelihood $Pr([x]|X)$ of $[x]$ with respect to X and the a priori probability $Pr(X)$. Similarly, for X^c we also have

$$Pr(X^c|[x]) = \frac{Pr([x]|X^c)Pr(X^c)}{Pr([x])} . \tag{24.40}$$

By substituting results of (24.39) and (24.40) into (24.38), we immediately have

$$\begin{aligned}\operatorname{logit}(Pr(X|[x])) &= \log O(X|[x]) \\ &= \log \frac{Pr(X|[x])}{Pr(X^c|[x])} \\ &= \log \frac{Pr([x]|X)}{Pr([x]|X^c)} \cdot \frac{Pr(X)}{Pr(X^c)} \\ &= \log \frac{Pr([x]|X)}{Pr([x]|X^c)} + \log O(X) ,\end{aligned} \tag{24.41}$$

where $O(X|[x])$ and $O(X)$ are the a posterior and the a prior odds, respectively, and $Pr([x]|X)/Pr([x]|X^c)$ is the likelihood ratio.

A threshold value on the probability can be expressed as another threshold value on logarithm of the likelihood ratio. For the positive region, we have

$$\begin{aligned}&Pr(X|[x]) \geq \alpha \\ &\iff \log \frac{Pr(X|[x])}{Pr(X^c|[x])} \geq \log \frac{\alpha}{1-\alpha} \\ &\iff \log \left(\frac{Pr([x]|X)}{Pr([x]|X^c)} \cdot \frac{Pr(X)}{Pr(X^c)}\right) \geq \log \frac{\alpha}{1-\alpha} \\ &\iff \log \frac{Pr([x]|X)}{Pr([x]|X^c)} \geq \log \frac{Pr(X^c)}{Pr(X)} + \log \frac{\alpha}{1-\alpha} \\ &= \alpha' .\end{aligned} \tag{24.42}$$

Similar expressions can be obtained for the negative and boundary regions. The three regions can now be written as

$$\begin{aligned}\mathrm{POS}_{(\alpha,\beta)}(X) &= \left\{x \in U \mid \log \frac{Pr([x]|X)}{Pr([x]|X^c)} \geq \alpha'\right\} , \\ \mathrm{BND}_{(\alpha,\beta)}(X) &= \left\{x \in U \mid \beta' < \log \frac{Pr([x]|X)}{Pr([x]|X^c)} < \alpha'\right\} , \\ \mathrm{NEG}_{(\alpha,\beta)}(X) &= \left\{x \in U \mid \log \frac{Pr([x]|X)}{Pr([x]|X^c)} \leq \beta'\right\} ,\end{aligned} \tag{24.43}$$

where

$$\begin{aligned}\alpha' &= \log \frac{Pr(X^c)}{Pr(X)} + \log \frac{\alpha}{1-\alpha} , \\ \beta' &= \log \frac{Pr(X^c)}{Pr(X)} + \log \frac{\beta}{1-\beta} .\end{aligned} \tag{24.44}$$

With the transformation, we need to estimate the likelihoods that are relatively easier to obtain.

Suppose that an equivalence relation E_A is defined by using a subset of attributes $A \subseteq AT$. In the naive Bayesian rough set model, we estimate the likelihood ratio $Pr([x]_A|X)/Pr([x]_A|X^c)$ through the likelihoods $Pr([x]_a|X)$ and $Pr([x]_a|X^c)$ defined by individual attributes, as the latter can be estimated more accurately. For this purpose, based on the results in (24.6), we make the following naive conditional independence assumptions

$$\begin{aligned}Pr([x]_A|X) &= Pr\left(\bigcap_{a\in A}[x]_a|X\right) = \prod_{a\in A} Pr\left([x]_a|X\right) , \\ Pr([x]_A|X^c) &= Pr\left(\bigcap_{a\in A}[x]_a|X^c\right) = \prod_{a\in A} Pr([x]_a|X^c) .\end{aligned} \tag{24.45}$$

By inserting them into (24.42) and assuming that $[x]$ is defined by a subset of attributes $A \subseteq AT$, namely, $[x] = [x]_A$, we have

$$\begin{aligned}&\log \frac{Pr([x]_A|X)}{Pr([x]_A|X^c)} \geq \alpha' \\ &\iff \log \frac{\prod_{a\in A} Pr([x]_a|X)}{\prod_{a\in A} Pr([x]_a|X^c)} \geq \alpha' \\ &\iff \sum_{a\in A} \log \frac{Pr([x]_a|X)}{Pr([x]_a|X^c)} \geq \alpha' .\end{aligned} \tag{24.46}$$

Similar conditions can be derived for negative and boundary regions. Finally, the three regions can be defined as

$$\begin{aligned}\mathrm{POS}_{(\alpha,\beta)}(X) &= \left\{x \in U \mid \sum_{a\in A} \log \frac{Pr([x]_a|X)}{Pr([x]_a|X^c)} \geq \alpha'\right\} ,\\ \mathrm{BND}_{(\alpha,\beta)}(X) &= \left\{x \in U \mid \beta' \right.\\ &\qquad \left. < \sum_{a\in A} \log \frac{Pr([x]_a|X)}{Pr([x]_a|X^c)} < \alpha'\right\} ,\\ \mathrm{NEG}_{(\alpha,\beta)}(C) &= \left\{x \in U \mid \sum_{a\in A} \log \frac{Pr([x]_a|X)}{Pr([x]_a|X^c)} \leq \beta'\right\} ,\end{aligned} \tag{24.47}$$

where

$$\begin{aligned}\alpha' &= \log \frac{Pr(X^c)}{Pr(X)} + \log \frac{\alpha}{1-\alpha} ,\\ \beta' &= \log \frac{Pr(X^c)}{Pr(X)} + \log \frac{\beta}{1-\beta} .\end{aligned} \tag{24.48}$$

We obtain a model in which we only need to estimate likelihoods of equivalence classes induced by individual attributes.

The likelihoods $Pr([x]_a|X)$ and $Pr([x]_a|X^c)$ may be simply estimated based on the following frequencies

$$\begin{aligned}Pr([x]_a|X) &= \frac{|[x]_a \cap X|}{|X|} ,\\ Pr([x]_a|X^c) &= \frac{|[x]_a \cap X^c|}{|X^c|} ,\end{aligned}$$

where $[x]_a = \{y \in U \mid I_a(y) = I_a(x)\}$. An equivalence class defined by a single attribute is usually large in comparison with an equivalence classes defined by a subset of attributes. Probability estimation based on the former may be more accurate than based on the latter.

Naive Bayesian rough sets provide only one of possible ways to estimate the conditional probability. Other estimation methods include logistic regress [24.46] and the maximum likelihood estimators [24.63].

24.5.3 Three-Way Decisions: Interpreting the Three Regions

A theory of three-way decisions [24.64] is motivated by the needs for interpreting the three regions [24.65–67] and moves beyond rough sets. The main results of three-way decisions can be found in two recent books edited by *Jia* et al. [24.68] and *Liu* et al. [24.69], respectively. We present an interpretation of rough set three regions based on the framework of three-way decisions.

In an information table, with respect to a subset of attributes $A \subseteq AT$, an object x induces a logic formula

$$\bigwedge_{a\in A} a = I_a(x) , \tag{24.49}$$

where $I_a(x) \in V_a$ and the atomic formula $a = I_a(x)$ indicates that the value of an object on attribute a is $I_a(x)$. An object y satisfies the formula if $I_a(y) = I_a(x)$ for all $a \in A$, that is

$$\left(y \models \bigwedge_{a\in A} a = I_a(x)\right) \Longleftrightarrow \forall a \in A \;\; (I_a(y) = I_a(x)). \tag{24.50}$$

With these notations, we are ready to interpret rough set in three regions.

From the three regions, we can construct three classes of rules for classifying an object, called the positive, negative, and boundary rules [24.58, 66, 67]. They are expressed in the following forms, for $y \in U$:

- Positive rule induced by an equivalence class $[x] \subseteq \mathrm{POS}_{(\alpha,\beta)}(X)$

$$\textit{if } y \models \bigwedge_{a\in A} a = I_a(x), \;\; \text{accept } y \in X$$

- Negative rule induced by an equivalence class $[x] \subseteq \mathrm{NEG}_{(\alpha,\beta)}(X)$

$$\textit{if } y \models \bigwedge_{a\in A} a = I_a(x), \;\; \text{reject } y \in X$$

- Boundary rule induced by an equivalence class $[x] \subseteq \mathrm{BND}_{(\alpha,\beta)}(X)$

$$\begin{aligned}&\textit{if } y \models \bigwedge_{a\in A} a = I_a(x), \;\; \text{neither accept}\\ &\qquad\qquad \text{nor reject } y \in X .\end{aligned}$$

The three types of rules have very different semantic interpretations as defined by their respective decisions. A positive rule allows us to *accept* an object y to be a member of X, because y has a higher probability of being in X due to the facts that $y \in [x]_A$ and $Pr(X|[x]_A) \geq \alpha$. A negative rule enables us to *reject* an object y to be a member of X, because y has lower probability of being in X due to the facts that $y \in [x]_A$ and $Pr(X|[x]_A) \leq \beta$. When the probability of y being in X is neither high nor low, a boundary rule makes a noncommitment decision. Although we explicitly give the class of boundary rules for convenience and completeness, we do not really need this class, once we have both classes of positive and negative rules. Whenever we can not accept nor reject an object to be a member of X, we choose a noncommitment decision.

Both actions of acceptance and rejection as associated with errors and costs. The error rate of a positive rule is given by $1 - Pr(X|[x])$, which, by definition of the three regions, is at or below $1 - \alpha$. The error rate of negative rule is given by $Pr(X|[x])$ and is at or below β. It becomes clear that the introduction of a noncommitment decision is to ensure both a low level of acceptance error and a low level of rejection error. According to the 3×2 table in Sect. 24.5.1, the cost a positive rule is $\lambda_{PP} Pr(X|[x]_A) + \lambda_{PN}(1 - Pr(X|[x]_A))$ and is bounded above by $\alpha\lambda_{PP} + (1-\alpha)\lambda_{PN}$. The cost a negative rule is $\lambda_{NP} Pr(X|[x]_A) + \lambda_{NN}(1 - Pr(X|[x]_A))$ and is bounded above by $\beta\lambda_{NP} + (1-\beta)\lambda_{NN}$. From view of cost, a noncommitment decision is preferred if its cost is less than an action of acceptance or rejection.

24.6 Dominance-Based Rough Set Approaches

Very often value sets V_a of some attributes $a \in AT$ are ordered in the sense that it is meaningful to consider a binary relation $\succsim_a$ on V_a such that for $x, y \in U$, $I_a(x) \succsim_a I_a(y)$ means that x possesses some property related to attribute a at least as much as y. In this case, it is natural to consider $\succsim_a$ as complete preorder on V_a, i. e., a transitive and strongly complete binary relation on V_a (let us remember that strong completeness means that for all $v_a, u_a \in V_a$ we have $v_a \succsim_a u_a$ or $u_a \succsim_a v_a$ and that this implies the reflexivity of $\succsim_a$). Observe that the binary relation $\succsim_a^U$ on U defined as $x \succsim_a^U y$ if $I_a(x) \succsim_a I_a(y)$ for all $x, y \in U$ is a complete preorder. The first type of properties considered in this perspective were preferences encountered in Multiple Criteria Decision Aiding (MCDA) (for a comprehensive collection of state of the art surveys see [24.70]), where for $x, y \in U$, $I_a(x) \succsim I_a(y)$ means x is at least as good as y with respect to attribute a that in this case is called *criterion*. If there are attributes $a \in AT$ related to some complete preorder $\succsim_a$, then the indiscernibility relation is unable to produce granules in U taking into account the order generated by $\succsim_a$. To do so, the indiscernibility relation has to be substituted by a new binary relation on U that, using a term coming from MCDA, is called *dominance relation*. Suppose, for simplicity, all attributes a from AT are criteria related to corresponding complete preorders $\succsim_a$.

We say that x *dominates* y with respect to $A \subseteq AT$ (shortly, x *A-dominates* y) denoted by $x \succsim_A^U y$, if $I_a(x) \succsim_a I(y)$ for all $a \in A$. Since $\succsim_a^U$ is a complete preorder on U for each $a \in AT$, $\succsim_A$ is a partial preoder on U, i. e. $\succsim_A$ is a reflexive and transitive binary relation on U.

For any $x \in U$ and for each nonempty $A \subseteq AT$, we can define a positive and a negative cone of dominance, denoted by $D_A^+(x)$ and $D_A^-(x)$, respectively,

$$D_A^+(x) = \{y \in U \mid y \succsim_A^U x\} \;, \qquad D_A^-(x) = \{y \in U \mid x \succsim_A^U y\} \;. \tag{24.51}$$

For simplicity, we also simply write $D^+(x)$ and $D^-(x)$ when no confusion arises.

Let us explain how the rough set concept has been generalized to the dominance-based rough set approach (DRSA) in order to enable granular computing with dominance cones (for more details, see Chap. 22, and [24.31–34, 71–74]).

For any $X \subseteq U$ we define upward lower and upper approximations $\underline{apr}^+(X)$ and $\overline{apr}^+(X)$, as well as downward lower and upper approximations $\underline{apr}^-(X)$ and $\overline{apr}^-(X)$, as follows

$$\begin{aligned} \underline{apr}^+(X) &= \{x \in U \mid D^+(x) \subseteq X\} \;, \\ \overline{apr}^+(X) &= \{x \in U \mid D^-(x) \cap X \neq \emptyset\} \;, \\ \underline{apr}^-(X) &= \{x \in U \mid D^-(x) \subseteq X\} \;, \\ \overline{apr}^-(X) &= \{x \in U \mid D^+(x) \cap X \neq \emptyset\} \;. \end{aligned} \tag{24.52}$$

For any $X \subseteq U$, using cones of dominance $D^+(x)$ and $D^-(x)$, we can define three upward pairwise disjoint positive, negative and boundary regions

$$\begin{aligned} \mathrm{POS}^+(X) &= \{x \in U \mid D^+(x) \subseteq X\} \;, \\ \mathrm{NEG}^+(X) &= \{x \in U \mid D^-(x) \cap X = \emptyset\} \;, \\ \mathrm{BND}^+(X) &= \{x \in U \mid D^+(x) \not\subseteq X \\ &\qquad \text{and } D^-(x) \cap X \neq \emptyset\} \;. \end{aligned} \tag{24.53}$$

Analogously, for any $X \subseteq U$, we can define three downward pairwise disjoint positive, negative, and boundary regions

$$\begin{aligned}
\mathrm{POS}^-(X) &= \{x \in U \mid D^-(x) \subseteq X\}\,,\\
\mathrm{NEG}^-(X) &= \{x \in U \mid D^+(x) \cap X = \emptyset\}\,,\\
\mathrm{BND}^-(X) &= \{x \in U \mid D^-(x) \not\subseteq X\\
&\qquad \text{and } D^+(x) \cap X \neq \emptyset\}\,. \qquad (24.54)
\end{aligned}$$

Observe that the following complementarity properties hold: For all $X \subseteq U$

$$\begin{aligned}
\mathrm{POS}^+(X) &= \mathrm{NEG}^-(U-X)\,,\\
\mathrm{POS}^-(X) &= \mathrm{NEG}^+(U-X)\,,\\
\mathrm{BND}^+(X) &= \mathrm{BND}^-(U-X)\,,\\
\mathrm{BND}^-(X) &= \mathrm{BND}^+(U-X)\,. \qquad (24.55)
\end{aligned}$$

For all $X \subseteq U$, the pair of upward approximations and three upward regions determine each others as follows

$$\begin{aligned}
\mathrm{POS}^+(X) &= \underline{apr}^+(X)\,,\\
\mathrm{NEG}^+(X) &= (\overline{apr}^+(X))^c\,,\\
\mathrm{BND}^+(X) &= \overline{apr}^+(X) - \underline{apr}^+(X)\,, \qquad (24.56)
\end{aligned}$$

and

$$\begin{aligned}
\underline{apr}^+(X) &= \mathrm{POS}^+(X)\,,\\
\overline{apr}^+(X) &= \mathrm{POS}^+(X) \cup \mathrm{BND}(X)\,. \qquad (24.57)
\end{aligned}$$

Analogously, for all $X \subseteq U$, the pair of downward approximations and three downward regions determine each others as follows

$$\begin{aligned}
\mathrm{POS}^-(X) &= \underline{apr}^-(X)\,,\\
\mathrm{NEG}^-(X) &= (\overline{apr}^-(X))^c\,,\\
\mathrm{BND}^-(X) &= \overline{apr}^-(X) - \underline{apr}^-(X)\,, \qquad (24.58)
\end{aligned}$$

and

$$\begin{aligned}
\underline{apr}^-(X) &= \mathrm{POS}^-(X)\,,\\
\overline{apr}^-(X) &= \mathrm{POS}^-(X) \cup \mathrm{BND}(X)\,. \qquad (24.59)
\end{aligned}$$

24.7 A Basic Model of Dominance-Based Probabilistic Rough Sets

DRSA considers only qualitative relationship between positive and negative cones $D^+(x)$ and $D^-(x)$, and a set X, namely, a positive or negative cone is a subset of the set or has a nonempty intersection with the set. This qualitative nature becomes clearer with a probabilistic interpretation. Suppose $Pr(X|D^+(x))$ denotes the conditional probability that an object is in X, given that the object is in $D^+(x)$, as well as $Pr(X|D^-(x))$ denotes the conditional probability that an object is in X, given that the object is in $D^-(x)$. The conditions for defining rough set three upward regions can be equivalently expressed as

$$\begin{aligned}
&D^+(x) \subseteq X \iff Pr(X|D^+(x)) \geq 1\,;\\
&D^-(x) \cap X = \emptyset \iff Pr(X|D^-(x)) \leq 0\,;\\
&D^+(x) \not\subseteq X \wedge D^-(x) \cap X \neq \emptyset\\
&\quad \iff Pr(X|D^+(x)) < 1\\
&\qquad \wedge Pr(X|D^-(x)) > 0\,. \qquad (24.60)
\end{aligned}$$

Analogously, the conditions for defining rough set three upward regions can be equivalently expressed as

$$\begin{aligned}
&D^-(x) \subseteq X \iff Pr(X|D^-(x)) \geq 1\,;\\
&D^+(x) \cap X = \emptyset \iff Pr(X|D^+(x)) \leq 0\,;\\
&D^-(x) \not\subseteq X \wedge D^+(x) \cap X \neq \emptyset\\
&\quad \iff Pr(X|D^-(x)) < 1 \wedge Pr(X|D^+(x)) > 0\,. \qquad (24.61)
\end{aligned}$$

By those conditions, DRSA upward and downward three regions can be equivalently expressed as

$$\begin{aligned}
\mathrm{POS}^+(X) &= \{x \in U \mid Pr(X|D^+(x)) \geq 1\}\,,\\
\mathrm{NEG}^+(X) &= \{x \in U \mid Pr(X|D^-(x)) \leq 0\}\,,\\
\mathrm{BND}^+(X) &= \{x \in U \mid Pr(X|D^+(x)) < 1\\
&\qquad \wedge Pr(X|D^-(x)) > 0\}\,,\\
\mathrm{POS}^-(X) &= \{x \in U \mid Pr(X|D^-(x)) \geq 1\}\,,\\
\mathrm{NEG}^-(X) &= \{x \in U \mid Pr(X|D^+(x)) \leq 0\}\,,\\
\mathrm{BND}^-(X) &= \{x \in U \mid Pr(X|D^-(x)) < 1\\
&\qquad \wedge Pr(X|D^+(x)) > 0\}\,. \qquad (24.62)
\end{aligned}$$

Observe that DRSA approximations use only the two extreme values, i. e., 1 and 0, of probability.

It is natural to generalize DRSA approximations by replacing 1 and 0 with some other values in the unit interval $[0,1]$. Given a pair of thresholds α, β with $0 \leq \beta < \alpha \leq 1$, the main results of probabilistic DRSA are the (α, β)-probabilistic regions defined by

$$\begin{aligned}
\mathrm{POS}^{+}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|D^{+}(x)) \geq \alpha\}\,,\\
\mathrm{NEG}^{+}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|D^{-}(x)) \leq \beta\}\,,\\
\mathrm{BND}^{+}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|D^{+}(x)) < \alpha\\
&\qquad \wedge Pr(X|D^{-}(x)) > \beta\}\,,\\
\mathrm{POS}^{-}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|D^{-}(x)) \geq \alpha\}\,,\\
\mathrm{NEG}^{-}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|D^{+}(x)) \leq \beta\}\,,\\
\mathrm{BND}^{-}_{(\alpha,\beta)}(X) &= \{x \in U \mid Pr(X|D^{-}(x)) < \alpha\\
&\qquad \wedge Pr(X|D^{+}(x)) > \beta\}\,. \qquad (24.63)
\end{aligned}$$

The DRSA rough set model is a special case in which $\alpha = 1$ and $\beta = 0$. In the case when $0 < \beta = \alpha < 1$, the three regions are given by

$$\begin{aligned}
\mathrm{POS}^{+}_{(\alpha,\alpha)}(X) &= \{x \in U \mid Pr(X|D^{+}(x)) \geq \alpha\}\,,\\
\mathrm{NEG}^{+}_{(\alpha,\alpha)}(X) &= \{x \in U \mid Pr(X|D^{-}(x)) \leq \alpha\}\,,\\
\mathrm{BND}^{-}_{(\alpha,\alpha)}(X) &= \{x \in U \mid Pr(X|D^{+}(x)) < \alpha\\
&\qquad \wedge Pr(X|D^{-}(x)) > \alpha\}\,,\\
\mathrm{POS}^{-}_{(\alpha,\alpha)}(X) &= \{x \in U \mid Pr(X|D^{-}(x)) \geq \alpha\}\,,\\
\mathrm{NEG}^{-}_{(\alpha,\alpha)}(X) &= \{x \in U \mid Pr(X|D^{+}(x)) \leq \alpha\}\,,\\
\mathrm{BND}^{-}_{(\alpha,\alpha)}(X) &= \{x \in U \mid Pr(X|D^{-}(x)) < \alpha\\
&\qquad \wedge Pr(X|D^{+}(x)) > \alpha\}\,. \qquad (24.64)
\end{aligned}$$

24.8 Variants of Probabilistic Dominance-Based Rough Set Approach

Several models generalizing dominance-based rough sets by incorporating probabilistic information can be considered.

24.8.1 Variable Consistency Dominance-Based Rough Sets

In a first version of variable consistency dominance-based rough sets [24.23] (see also [24.24]) the standard set inclusions $D^{+}(x) \subseteq X$ and $D^{-}(x) \subseteq X$ can be generalized into graded set inclusion $s^{+}(D^{+}(x), X)$ and $s^{-}(D^{-}(x), X)$ called measure of the relative upward and downward degree of misclassification of $D^{+}(x)$ and $D^{-}(x)$ with respect to X, respectively. A particular upward and downward measure is given by

$$\begin{aligned}
s^{+}(D^{+}(x), X) &= 1 - \frac{|D^{+}(x) \cap X|}{|D^{+}(x)|}\,,\\
s^{-}(D^{-}(x), X) &= 1 - \frac{|D^{-}(x) \cap X|}{|D^{-}(x)|}\,.
\end{aligned} \qquad (24.65)$$

By introducing a threshold $0 \leq z < 0.5$, one can define three upward and downward regions as follows

$$\begin{aligned}
\mathrm{VPOS}^{+}_{z}(X) &= \{x \in U \mid s^{+}(D^{+}(x), X) \leq z\}\,,\\
\mathrm{VNEG}^{+}_{z}(X) &= \{x \in U \mid s^{-}(D^{-}(x), X) \geq 1 - z\}\,,\\
\mathrm{VBND}^{+}_{z}(X) &= \{x \in U \mid s^{+}(D^{+}(x), X) > z\\
&\qquad \wedge s^{-}(D^{-}(x), X) < 1 - z\}\,,\\
\mathrm{VPOS}^{-}_{z}(X) &= \{x \in U \mid s^{-}(D^{-}(x), X) \leq z\}\,,\\
\mathrm{VNEG}^{-}_{z}(X) &= \{x \in U \mid s^{+}(D^{+}(x), X) \geq 1 - z\}\,,\\
\mathrm{VBND}^{-}_{z}(X) &= \{x \in U \mid s^{-}(D^{-}(x), X) > z\\
&\qquad \wedge s^{+}(D^{+}(x), X) < 1 - z\}\,. \qquad (24.66)
\end{aligned}$$

A more generalized version using a pair of thresholds can be defined as follows: for $0 \leq l < u \leq 1$,

$$\begin{aligned}
\mathrm{VPOS}^{+}_{(l,u)}(X) &= \{x \in U \mid s^{+}(D^{+}(x), X) \leq l\}\,,\\
\mathrm{VNEG}^{+}_{(l,u)}(X) &= \{x \in U \mid s^{-}(D^{-}(x), X) \geq u\}\,,\\
\mathrm{VBND}^{+}_{(l,u)}(X) &= \{x \in U \mid s^{+}(D^{+}(x), X) > l\\
&\qquad \wedge s^{-}(D^{-}(x), X) < u\}\,,\\
\mathrm{VPOS}^{-}_{(l,u)}(X) &= \{x \in U \mid s^{-}(D^{-}(x), X) \leq l\}\,,\\
\mathrm{VNEG}^{-}_{(l,u)}(X) &= \{x \in U \mid s^{+}(D^{+}(x), X) \geq u\}\,,\\
\mathrm{VBND}^{-}_{(l,u)}(X) &= \{x \in U \mid s^{-}(D^{-}(x), X) > l\\
&\qquad \wedge s^{+}(D^{+}(x), X) < u\}\,. \qquad (24.67)
\end{aligned}$$

The one-threshold model may be considered as a special case of the two-threshold model with $l = z$ and $u = 1 - z$.

One may interpret the ratio in (24.65) as an estimation of the conditional probability $Pr(X|D^+(x))$ and $Pr(X|D^-(x))$, namely,

$$\begin{aligned}
s^+(Pr(X|D^+(x)), X) &= 1 - \frac{|D^+(x) \cap X|}{|D^+(x)|} \\
&= 1 - Pr(X|D^+(x)) \ , \\
s^-(Pr(X|D^-(x)), X) &= 1 - \frac{|D^-(x) \cap X|}{|D^-(x)|} \\
&= 1 - Pr(X|D^-(x)) \ .
\end{aligned} \tag{24.68}$$

By setting $\alpha = 1 - l$ and $\beta = 1 - u$, we immediately get

$$\begin{aligned}
\mathrm{POS}^+_{A,(1-l,1-u)}(X) &= \{x \in U \mid Pr(X|D^+(x)) \geq 1 - l\} \\
&= \{x \in U \mid s^+(D^+(x), X) \leq l\} \\
&= \mathrm{VPOS}^+_{(l,u)}(X) \ , \\
\mathrm{NEG}^+_{(1-l,1-u)}(X) &= \{x \in U \mid Pr(X|D^-(x)) \leq 1 - u\} \\
&= \{x \in U \mid s^-(D^-(x), X) \geq u\} \\
&= \mathrm{VNEG}^+_{(l,u)}(X) \ , \\
\mathrm{BND}^+_{(1-l,1-u)}(X) &= \{x \in U \mid Pr(X|D^+(x)) < 1 - l \\
&\quad \wedge Pr(X|D^-(x)) > 1 - u\} \\
&= \{x \in U \mid s^+(D^+(x), X) > l \\
&\quad \wedge s^-(D^-(x), X) < u\} \\
&= \mathrm{VBND}^+_{(l,u)}(X) \ , \\
\mathrm{POS}^-_{(1-l,1-u)}(X) &= \{x \in U \mid Pr(X|D^-(x)) \geq 1 - l\} \\
&= \{x \in U \mid s^-(D^-(x), X) \leq l\} \\
&= \mathrm{VPOS}^-_{(l,u)}(X) \ , \\
\mathrm{NEG}^-_{(1-l,1-u)}(X) &= \{x \in U \mid Pr(X|D^+(x)) \leq 1 - u\} \\
&= \{x \in U \mid s^+(D^+(x), X) \geq u\} \\
&= \mathrm{VNEG}^+_{(l,u)}(X) \ , \\
\mathrm{BND}^+_{(1-l,1-u)}(X) &= \{x \in U \mid Pr(X|D^-(x)) < 1 - l \\
&\quad \wedge Pr(X|D^+(x)) > 1 - u\} \\
&= \{x \in U \mid s^-(D^-(x), X) > l \\
&\quad \wedge s^+(D^+(x), X) < u\} \\
&= \mathrm{VBND}^-_{(l,u)}(X) \ .
\end{aligned} \tag{24.69}$$

24.8.2 Parameterized Dominance-Based Rough Sets

Parameterized rough sets based on dominance [24.24] generalize variable consistency DRSA by introducing a Bayesian confirmation measure and a pair of thresholds on the confirmation measure, in addition to a pair of thresholds on conditional probability. Let $c^+(D^+(x), X)$ and $c^-(D^-(x), X)$ denote a Bayesian upward and downward confirmation measure, respectively, that indicate the degree to which positive or negative cones $D^+(x)$ and $D^-(x)$ confirm the hypothesis X. The upward and downward Bayesian confirmation measures corresponding to those ones introduced in Sect. 24.4.3 are

$$\begin{aligned}
c^+_d(D^+(x), X) &= Pr(X|D^+(x)) - Pr(X) \ , \\
c^-_d(D^-(x), X) &= Pr(X|D^-(x)) - Pr(X) \ , \\
c^+_r(D^+(x), X) &= \frac{Pr(X|D^+(x))}{Pr(X)} \ , \\
c^-_r(D^+(x), X) &= \frac{Pr(X|D^-(x))}{Pr(X)} \ .
\end{aligned} \tag{24.70}$$

Given a pair of thresholds (s, t) with $t < s$, three (α, β, s, t)-parameterized regions can be defined as follows

$$\begin{aligned}
\mathrm{PPOS}^+_{(\alpha,\beta,s,t)}(X) &= \{x \in U \mid Pr(X|D^+(x)) \geq \alpha \\
&\quad \wedge c^+(D^+(x), X) \geq s\} \ , \\
\mathrm{PNEG}^+_{(\alpha,\beta,s,t)}(X) &= \{x \in U \mid Pr(X|D^-(x)) \leq \beta \\
&\quad \wedge c^-(D^-(x), X) \leq t\} \ , \\
\mathrm{PBND}^+_{(\alpha,\beta,s,t)}(X) &= \{x \in U \mid (Pr(X|D^+(x)) < \alpha \\
&\quad \vee c^+(D^+(x), X) < s) \\
&\quad \wedge (Pr(X|D^-(x)) > \beta \\
&\quad \vee c^-(D^-(x), X) > t)\} \\
\mathrm{PPOS}^-_{(\alpha,\beta,s,t)}(X) &= \{x \in U \mid Pr(X|D^-(x)) \geq \alpha \\
&\quad \wedge c^-(D^-(x), X) \geq s\} \ , \\
\mathrm{PNEG}^-_{(\alpha,\beta,s,t)}(X) &= \{x \in U \mid Pr(X|D^+(x)) \leq \beta \\
&\quad \wedge c^+(D^+(x), X) \leq t\} \ , \\
\mathrm{PBND}^-_{(\alpha,\beta,s,t)}(X) &= \{x \in U \mid (Pr(X|D^-(x)) < \alpha \\
&\quad \vee c^-(D^-(x), X) < s) \\
&\quad \wedge (Pr(X|D^+(x)) > \beta \\
&\quad \vee c^+(D^+(x), X) > t)\} \ .
\end{aligned} \tag{24.71}$$

Let us remember that a family of consistency measures, called gain-type consistency measures, and inconsistency measures, called cost-type consistency measures, larger than confirmation measures, and the related dominance-based rough sets have been considered in [24.24]. For any $x \in U$ and $X \subseteq U$, for a consistency measure $m_c(x,X)$, x can be assigned to the positive region of X if $m_c(x,X) \geq \alpha$, with α being a proper threshold, while for an inconsistency measure $m_{ic}(x,X)$, x can be assigned to the positive region of X if $m_c(x,X) \leq \alpha$. A consistency measure $m_c(x,X)$ or an inconsistency measure $m_{ic}(x,X)$ are monotonic (Sect. 22.3.2 in Chap. 22) if they do not deteriorate when:

(m1) The set of attributes is growing,
(m2) The set of objects is growing,
(m3) x improves its evaluation, so that it dominates more objects.

Among the considered consistency and inconsistency measures, one that can be considered very interesting because it enjoys all the considered monotonity properties (m1)–(m3) while maintaining a reasonably easy formulation is the inconsistency measures ε' which is expressed as follows:

- In the case of dominance-based upward approximation

$$\varepsilon'^{+}(x,X) = \frac{|D^{+}(x) \cap (U-X)|}{|X|}$$

- In the case of dominance-based downward approximation

$$\varepsilon'^{-}(x,X) = \frac{|D^{-}(x) \cap (U-X)|}{|X|} .$$

Observe that as explained in [24.24], consistency and inconsistency measures can be properly reformulated in order to be used in indiscernbility-based rough sets. For example, inconsistency measure ε' in case of indicernibility-based rough sets becomes

$$\varepsilon'^{+}(x,X) = \frac{|[x] \cap (U-X)|}{|X|} .$$

24.8.3 Confirmation-Theoretic Dominance-Based Rough Sets

A separation of the parameterized model into two models within DRSA can be constructed as follows. One is the conventional probabilistic model and the other is a confirmation-theoretic model. For an upward and a downward Bayesian confirmation measure $(c^{+}(D^{+}(x),X)$ and $c^{-}(D^{-}(x),X))$, and a pair of thresholds (s,t) with $t < s$, three confirmation regions are defined by

$$\begin{aligned}
\mathrm{CPOS}^{+}_{(s,t)}(X) &= \{x \in U \mid c^{+}(D^{+}(x),X) \geq s\} ,\\
\mathrm{CNEG}^{+}_{(s,t)}(X) &= \{x \in U \mid c^{-}(D^{-}(x),X) \leq t\} ,\\
\mathrm{CBND}^{+}_{(s,t)}(X) &= \{x \in U \mid c^{+}(D^{+}(x),X) < s \\
&\qquad \wedge c^{-}(D^{-}(x),X) > t\} ,\\
\mathrm{CPOS}^{-}_{(s,t)}(X) &= \{x \in U \mid c^{+}(D^{-}(x),X) \geq s\} ,\\
\mathrm{CNEG}^{-}_{(s,t)}(X) &= \{x \in U \mid c^{-}(D^{+}(x),X) \leq t\} ,\\
\mathrm{CBND}^{-}_{(s,t)}(X) &= \{x \in U \mid c^{-}(D^{-}(x),X) < s \\
&\qquad \wedge c^{+}(D^{+}(x),X) > t\} .
\end{aligned} \tag{24.72}$$

For the case with $s = t$, we define

$$\begin{aligned}
\mathrm{CPOS}^{+}_{(s,s)}(X) &= \{x \in U \mid c^{+}(D^{+}(x),X) \geq s\} ,\\
\mathrm{CNEG}^{+}_{(s,s)}(X) &= \{x \in U \mid c^{-}(D^{-}(x),X) \leq s\} ,\\
\mathrm{CBND}^{+}_{(s,s)}(X) &= \{x \in U \mid c^{+}(D^{+}(x),X) < s \\
&\qquad \wedge c^{-}(D^{-}(x),X) > s\} ,\\
\mathrm{CPOS}^{-}_{(s,s)}(X) &= \{x \in U \mid c^{+}(D^{+}(x),X) \geq s\} ,\\
\mathrm{CNEG}^{-}_{(s,s)}(X) &= \{x \in U \mid c^{-}(D^{-}(x),X) \leq s\} ,\\
\mathrm{CBND}^{-}_{(s,s)}(X) &= \{x \in U \mid c^{+}(D^{+}(x),X) < s \\
&\qquad \wedge c^{+}(D^{+}(x),X) > s\} .
\end{aligned} \tag{24.73}$$

24.8.4 Bayesian Dominance-Based Rough Sets

Bayesian DRSAmodel in which the required pair of thresholds is interpreted using a priori probability $Pr(X)$ can be defined as an extension of the Bayesian DRSA and variable consistency Bayesian DRSA, as explained below.

For the Bayesian DRSA, the three upward and downward regions are defined by

$$\begin{aligned}
\mathrm{BPOS}^{+}(X) &= \{x \in U \mid Pr(X|D^{+}(x)) > Pr(X)\} ,\\
\mathrm{BNEG}^{+}(X) &= \{x \in U \mid Pr(X|D^{-}(x)) < Pr(X)\} ,\\
\mathrm{BBND}^{+}(X) &= \{x \in U \mid Pr(X|D^{+}(x)) \leq Pr(X) \\
&\qquad \wedge Pr(X|D^{-}(x)) \geq Pr(X)\} ,\\
\mathrm{BPOS}^{-}(X) &= \{x \in U \mid Pr(X|D^{-}(x)) > Pr(X)\} ,\\
\mathrm{BNEG}^{-}(X) &= \{x \in U \mid Pr(X|D^{+}(x)) < Pr(X)\} ,\\
\mathrm{BBND}^{-}(X) &= \{x \in U \mid Pr(X|D^{-}(x)) \leq Pr(X) \\
&\qquad \wedge Pr(X|D^{+}(x)) \geq Pr(X)\} .
\end{aligned} \tag{24.74}$$

Bayesian dominance-based rough sets can be viewed as a special case of the decision-theoretic DRSA when $\alpha = \beta = Pr(X)$.

Recalling the upward and downward DRSA, Bayesian confirmation measures

$$c_d^+(D^+(x), X) = Pr(X|D^+(x)) - Pr(X)$$

and

$$c_d^-(D^-(x), X) = Pr(X|D^-(x)) - Pr(X) ,$$

Bayesian dominance-based rough sets can be expressed as confirmation-theoretic dominance-based rough sets as follows

$$\begin{aligned}
\mathrm{BPOS}^+(X) &= \{x \in U \mid Pr(X|D^+(x)) > Pr(X)\} ,\\
&= \{x \in U \mid c_d^+(D^+(x), X) > 0\} ,\\
\mathrm{BNEG}^+(X) &= \{x \in U \mid Pr(X|D^-(x)) < Pr(X)\} ,\\
&= \{x \in U \mid c_d^-(D^-(x), X) < 0\} ,\\
\mathrm{BBND}(X)^+ &= \{x \in U \mid Pr(X|D^+(x)) \leq Pr(X)\\
&\quad \wedge Pr(X|D^-(x)) \geq Pr(X)\}\\
&= \{x \in U \mid c_d^+(D^+(x), X) \leq 0\\
&\quad \wedge c_d^-(D^-(x), X) \geq 0\} ,\\
\mathrm{BPOS}^-(X) &= \{x \in U \mid Pr(X|D^-(x)) > Pr(X)\} ,\\
&= \{x \in U \mid c_d^-(D^-(x), X) > 0\} ,\\
\mathrm{BNEG}^+(X) &= \{x \in U \mid Pr(X|D^-(x)) < Pr(X)\} ,\\
&= \{x \in U \mid c_d^+(D^+(x), X) < 0\} ,\\
\mathrm{BBND}(X)^+ &= \{x \in U \mid Pr(X|D^-(x)) \leq Pr(X)\\
&\quad \wedge Pr(X|D^+(x)) \geq Pr(X)\}\\
&= \{x \in U \mid c_d^-(D^-(x), X) \leq 0\\
&\quad \wedge c_d^+(D^+(x), X) \geq 0\} .
\end{aligned} \tag{24.75}$$

That is, the Bayesian rough sets are models of confirmation-theoretic rough sets characterized by the upward and downward Bayesian confirmation measures c_d^+ and c_d^- with a pair of thresholds $s = t = 0$.

The three upward and downward regions of the variable precision Bayesian rough sets are defined as follows: for $\epsilon \in [0, 1)$,

$$\begin{aligned}
\mathrm{VBPOS}_\epsilon^+(X) &= \{x \in U \mid Pr(X|[x])\\
&\quad \geq 1 - \epsilon(1 - Pr(X))\} ,\\
\mathrm{VBNEG}_\epsilon(X) &= \{x \in U \mid Pr(X|[x]) \leq \epsilon Pr(X)\} ,\\
\mathrm{VBBND}_\epsilon(X) &= \{x \in U \mid \epsilon Pr(X) < Pr(X|[x])\\
&\quad < 1 - \epsilon(1 - Pr(X))\} .
\end{aligned} \tag{24.76}$$

24.9 Three Fundamental Issues of Probabilistic Dominance-Based Rough Sets

Also for probabilistic dominance-based rough sets, one must consider the three fundamental issues of interpretation and determination of the required pair of thresholds, estimation of the required conditional probabilities, and interpretation and applications of three probabilistic regions.

These three issues are considered in this section with respect to dominance-based rough sets.

24.9.1 Decision-Theoretic Dominance-Based Rough Set Model: Determining the Thresholds

Following [24.75], a decision-theoretic model formulates the construction of dominance-based rough set approximations as a Bayesian decision problem with a set of two states $\Omega = \{X, X^c\}$, indicating that an element is in X and not in X, respectively. In the case of upward dominance-based rough sets, we consider a set of three actions $\mathcal{A}^+ = \{a_P^+, a_B^+, a_N^+\}$, with a_P^+ deciding $x \in \mathrm{POS}^+(X)$, a_B^+ deciding $x \in \mathrm{BND}^+(X)$, and a_N^+ deciding $x \in \mathrm{NEG}^+(X)$, respectively. In case of downward dominance-based rough sets, we consider a set of three actions $\mathcal{A}^- = \{a_P^-, a_B^-, a_N^-\}$, with a_P^- deciding $x \in \mathrm{POS}^-(X)$, a_B^- deciding $x \in \mathrm{BND}^-(X)$, and a_N^- deciding $x \in \mathrm{NEG}^-(X)$, respectively. The losses regarding the actions for different states are given by the 6×2 matrix

	X (P)	X^c (N)
a_P^+	λ_{PP}^+	λ_{PN}^+
a_B^+	λ_{BP}^+	λ_{BN}^+
a_N^+	λ_{NP}^+	λ_{NN}^+
a_P^-	λ_{PP}^-	λ_{PN}^-
a_B^-	λ_{BP}^-	λ_{BN}^-
a_N^-	λ_{NP}^-	λ_{NN}^-

In the matrix:

- In the case that upward dominance-based rough approximations are considered, λ^+_{PP}, λ^+_{BP}, and λ^+_{NP} denote the losses incurred for taking actions a^+_P, a^+_B, and a^+_N, respectively, when an object belongs to X, and λ^+_{PN}, λ^+_{BN} and λ^+_{NN} denote the losses incurred for taking the same actions when the object does not belong to X,
- In the case that downward dominance-based rough approximations are considered, λ^-_{PP}, λ^-_{BP}, and λ^-_{NP} denote the losses incurred for taking actions a^-_P, a^-_B, and a^-_N, respectively, when an object belongs to X, and λ^-_{PN}, λ^-_{BN} and λ^-_{NN} denote the losses incurred for taking the same actions when the object does not belong to X.

In the case that upward dominance-based rough approximations are considered, the expected losses associated with taking different actions for objects in $D^+(x)$ can be expressed as

$$\begin{aligned} R(a^+_P|D^+(x)) &= \lambda^+_{PP}Pr(X|D^+(x)) \\ &\quad + \lambda^+_{PN}Pr(X^c|D^+(x)) , \\ R(a^+_B|D^+(x)) &= \lambda^+_{BP}Pr(X|D^+(x)) \\ &\quad + \lambda^+_{BN}Pr(X^c|D^+(x)) , \\ R(a^+_N|D^+(x)) &= \lambda^+_{NP}Pr(X|D^+(x)) \\ &\quad + \lambda^+_{NN}Pr(X^c|D^+(x)) . \end{aligned} \tag{24.77}$$

In the case that downward dominance-based rough approximations are considered, the expected losses associated with taking different actions for objects in $D^-(x)$ can be expressed as

$$\begin{aligned} R(a^-_P|D^-(x)) &= \lambda^-_{PP}Pr(X|D^-(x)) \\ &\quad + \lambda^-_{PN}Pr(X^c|D^-(x)) , \\ R(a^-_B|D^-(x)) &= \lambda^-_{BP}Pr(X|D^-(x)) \\ &\quad + \lambda^-_{BN}Pr(X^c|D^-(x)) , \\ R(a^-_N|D^-(x)) &= \lambda^-_{NP}Pr(X|D^-(x)) \\ &\quad + \lambda^-_{NN}Pr(X^c|D^-(x)) . \end{aligned} \tag{24.78}$$

In the case that upward dominance-based rough approximations are considered, the Bayesian decision procedure suggests the following minimum-risk decision rules

(P$^+$) If $R(a^+_P|D^+(x)) \le R(a^+_B|D^+(x))$
and $R(a^+_P|[x]) \le R(a^+_N|D^+(x))$,
decide $x \in \mathrm{POS}^+(X)$;

(B$^+$) If $R(a^+_B|D^+(x)) \le R(a^+_P|D^+(x))$
and $R(a^+_B|[x]) \le R(a^+_N|D^+(x))$,
decide $x \in \mathrm{BND}^+(X)$;

(N$^+$) If $R(a^+_N|D^+(x)) \le R(a^+_P|D^+(x))$
and $R(a^+_N|[x]) \le R(a^+_B|D^+(x))$,
decide $x \in \mathrm{NEG}^+(X)$.

In the case that downward dominance-based rough approximations are considered, the Bayesian decision procedure suggests the following minimum-risk decision rules

(P$^-$) If $R(a^-_P|D^-(x)) \le R(a^-_B|D^-(x))$
and $R(a^-_P|[x]) \le R(a^-_N|D^-(x))$,
decide $x \in \mathrm{POS}^-(X)$;

(B$^-$) If $R(a^-_B|D^-(x)) \le R(a^-_P|D^-(x))$
and $R(a^-_B|[x]) \le R(a^-_N|D^-(x))$,
decide $x \in \mathrm{BND}^-(X)$;

(N$^-$) If $R(a^-_N|D^-(x)) \le R(a^-_P|D^-(x))$
and $R(a^-_N|[x]) \le R(a^-_B|D^-(x))$,
decide $x \in \mathrm{NEG}^-(X)$.

Also in the case that dominance-based rough approximations are considered, when two or three actions have the same risk, one can use the same ordering for breaking a tie used in case indiscernibility-based rough approximations are used: a^+_P, a^+_N, a^+_B in case upward rough approximations are considered, and a^-_P, a^-_N, a^-_B in case downward rough approximations are considered.

Analogously to Sect. 24.5.1, let us consider the special class of loss functions with

$$\begin{aligned} &(\mathrm{c0}^+).\ \lambda^+_{PP} \le \lambda^+_{BP} < \lambda^+_{NP}, \quad \lambda^+_{NN} \le \lambda^+_{BN} < \lambda^+_{PN} , \\ &(\mathrm{c0}^-).\ \lambda^-_{PP} \le \lambda^-_{BP} < \lambda^-_{NP}, \quad \lambda^-_{NN} \le \lambda^-_{BN} < \lambda^-_{PN} . \end{aligned} \tag{24.79}$$

With the conditions (c0$^+$) and (c0$^-$), and the equations

$$Pr(X|D^+(x)) + Pr(X^c|D^+(x)) = 1$$

and

$$Pr(X|D^-(x)) + Pr(X^c|D^-(x)) = 1 ,$$

we can express the decision rules (P$^+$)–(N$^+$) and (P$^-$)–(N$^-$) in the following simplified form

(P$^+$) If $Pr(X|D^+(x)) \geq \alpha^+$
and $Pr(X|D^+(x)) \geq \gamma^+$,
decide $x \in \mathrm{POS}^+(X)$;

(B$^+$) If $Pr(X|D^+(x)) \leq \alpha^+$
and $Pr(X|D^+(x)) \geq \beta^+$,
decide $x \in \mathrm{BND}^+(X)$;

(N$^+$) If $Pr(X|D^+(x)) \leq \beta^+$
and $Pr(X|D^+(x)) \leq \gamma^+$,
decide $x \in \mathrm{NEG}^+(X)$;

(P$^-$) If $Pr(X|D^-(x)) \geq \alpha^-$
and $Pr(X|D^-(x)) \geq \gamma^+$,
decide $x \in \mathrm{POS}^-(X)$;

(B$^-$) If $Pr(X|D^-(x)) \leq \alpha^-$
and $Pr(X|D^-(x)) \geq \beta^+$,
decide $x \in \mathrm{BND}^-(X)$;

(N$^-$) If $Pr(X|D^-(x)) \leq \beta^-$
and $Pr(X|D^-(x)) \leq \gamma^+$,
decide $x \in \mathrm{NEG}^-(X)$.

where

$$
\begin{aligned}
\alpha^+ &= \frac{(\lambda^+_{PN} - \lambda^+_{BN})}{(\lambda^+_{PN} - \lambda^+_{BN}) + (\lambda^+_{BP} - \lambda^+_{PP})} ,\\
\beta^+ &= \frac{(\lambda^+_{BN} - \lambda^+_{NN})}{(\lambda^+_{BN} - \lambda^+_{NN}) + (\lambda^+_{NP} - \lambda^+_{BP})} ,\\
\gamma^+ &= \frac{(\lambda^+_{PN} - \lambda^+_{NN})}{(\lambda^+_{PN} - \lambda^+_{NN}) + (\lambda^+_{NP} - \lambda^+_{PP})} ,\\
\alpha^- &= \frac{(\lambda^-_{PN} - \lambda^-_{BN})}{(\lambda^-_{PN} - \lambda^-_{BN}) + (\lambda^-_{BP} - \lambda^-_{PP})} ,\\
\beta^- &= \frac{(\lambda^-_{BN} - \lambda^-_{NN})}{(\lambda^-_{BN} - \lambda^-_{NN}) + (\lambda^-_{NP} - \lambda^-_{BP})} ,\\
\gamma^- &= \frac{(\lambda^-_{PN} - \lambda^-_{NN})}{(\lambda^-_{PN} - \lambda^-_{NN}) + (\lambda^-_{NP} - \lambda^-_{PP})} .
\end{aligned}
\tag{24.80}
$$

By setting $\alpha^+ > \beta^+$ and $\alpha^- > \beta^-$, we obtain that $1 \geq \alpha^+ > \gamma^+ > \beta^+ \geq 0$ and $1 \geq \alpha^- > \gamma^- > \beta^- \geq 0$, that, after tie breaking, we give the following simplified rules

(P$^+$) If $Pr(X|D^+(x)) \geq \alpha^+$,
decide $x \in \mathrm{POS}^+(X)$;

(B$^+$) If $\beta^+ < Pr(X|D^+(x)) < \alpha^+$,
decide $x \in \mathrm{BND}^+(X)$;

(N$^+$) If $Pr(X|D^+(x)) \leq \beta^+$,
decide $x \in \mathrm{NEG}^+(X)$,

(P$^-$) If $Pr(X|D^-(x)) \geq \alpha^-$,
decide $x \in \mathrm{POS}^-(X)$;

(B$^-$) If $\beta^- < Pr(X|D^-(x)) < \alpha^-$,
decide $x \in \mathrm{BND}^-(X)$;

(N$^-$) If $Pr(X|D^-(x)) \leq \beta^-$,
decide $x \in \mathrm{NEG}^-(X)$,

so that the parameters γ^+ and γ^- are no longer needed. Each object can be put into one and only one upward region, and one and only one downward region by using rules (P$^+$), (B$^+$) and (N$^+$), and (P$^-$), (B$^-$) and (N$^-$), respectively. The upward (α^+, β^+)-probabilistic positive, negative, and boundary regions and downward (α^-, β^-)-probabilistic positive, negative, and boundary regions are given, respectively, by

$$
\begin{aligned}
\mathrm{POS}^+_{(\alpha^+,\beta^+)}(X) &= \{x \in U \mid Pr(X|D^+(x)) \geq \alpha^+\} ,\\
\mathrm{BND}^+_{(\alpha^+,\beta^+)}(X) &= \{x \in U \mid \beta^+ < Pr(X|D^+(x)) \\
&\qquad < \alpha^+\} ,\\
\mathrm{NEG}^+_{(\alpha^+,\beta^+)}(X) &= \{x \in U \mid Pr(X|D^+(x)) \leq \beta\} ,\\
\mathrm{POS}^-_{(\alpha^-,\beta^-)}(X) &= \{x \in U \mid Pr(X|D^-(x)) \geq \alpha^-\} ,\\
\mathrm{BND}^-_{(\alpha^-,\beta^-)}(X) &= \{x \in U \mid \beta^- < Pr(X|D^-(x)) \\
&\qquad < \alpha^-\} ,\\
\mathrm{NEG}_{(\alpha^-,\beta^-)}(X) &= \{x \in U \mid Pr(X|D^-(x)) \leq \beta^-\} .
\end{aligned}
\tag{24.81}
$$

An alternative decision theoretic model for dominance-based rough sets taking into account in the cost function the conditional probabilities $P(X|D^+(x))$ and $P(X^c|D^-(x))$ for upward rough approximations, as well as the conditional probabilities $P(X|D^-(x))$ and $P(X^c|D^+(x))$ for downward rough approximations, can be defined as follows.

In the case that upward dominance-based rough approximations are considered, the expected losses associated with taking different actions for objects in

$D^+(x)$ can be expressed as

$$\begin{aligned} R(a_P^+|D^+(x), D^-(x)) &= \lambda_{PP}^+ Pr(X|D^+(x)) \\ &\quad + \lambda_{PN}^+ Pr(X^c|D^-(x)) , \\ R(a_B^+|D^+(x), D^-(x)) &= \lambda_{BP}^+ Pr(X|D^+(x)) \\ &\quad + \lambda_{BN}^+ Pr(X^c|D^-(x)) , \\ R(a_N^+|D^+(x), D^-(x)) &= \lambda_{NP}^+ Pr(X|D^+(x)) \\ &\quad + \lambda_{NN}^+ Pr(X^c|D^-(x)) . \end{aligned} \tag{24.82}$$

In the case that downward dominance-based rough approximations are considered, the expected losses associated with taking different actions for objects in $D^-(x)$ can be expressed as

$$\begin{aligned} R(a_P^-|D^+(x), D^-(x)) &= \lambda_{PP}^- Pr(X|D^-(x)) \\ &\quad + \lambda_{PN}^- Pr(X^c|D^+(x)) , \\ R(a_B^-|D^+(x), D^-(x)) &= \lambda_{BP}^- Pr(X|D^-(x)) \\ &\quad + \lambda_{BN}^- Pr(X^c|D^+(x)) , \\ R(a_N^-|D^+(x), D^-(x)) &= \lambda_{NP}^- Pr(X|D^-(x)) \\ &\quad + \lambda_{NN}^- Pr(X^c|D^+(x)) . \end{aligned} \tag{24.83}$$

24.9.2 Stochastic Dominance-Based Rough Set Approach: Estimating the Conditional Probability

Naive Bayesian rough set model presented for rough sets based on indiscernibility in Sect. 24.5.2 can be extended quite straightforwardly to rough sets based on dominance. Thus, in this section, we present a different approach to estimate probabilities for rough approximations: stochastic rough set approach [24.25] (see also Sect. 22.3.3 in Chap. 22). It can be applied also to rough sets based on indiscernibility, but here we present this approach taking into consideration rough sets based on dominance. In the following, we shall consider upward dominance-based approximations of a given $X \subseteq U$. However, the same approach can be used for downward dominance-based approximations. From a probabilistic point of view, the assignment of object x to $X \subseteq U$ can be made with probability $Pr(X|D^+(x))$ and $Pr(X|D^-(x))$. This probability is supposed to satisfy the usual axioms of probability

$$\begin{aligned} &Pr(U|D^+(x)) = 1 , \\ &\quad Pr(U-X|D^+(x)) = 1 - Pr(X|D^+(x)) , \\ &Pr(U|D^-(x)) = 1 , \\ &\quad Pr(U-X|D^-(x)) = 1 - Pr(X|D^-(x)) . \end{aligned}$$

Moreover, this probability has to satisfy an axiom related to the choice of the rough upward approximation, i.e., the positive monotonic relationships one expects between membership in $X \subseteq U$ and possession of the properties related to attributes from AT, i.e., the dominance relation $\succsim$: for any $x, y \in U$ such that $x \succsim y$

$$\begin{aligned} &(i)\ Pr(X|D^+(x)) \geq Pr(X|D^+(y)) , \\ &(ii)\ Pr(U-X|D^-(x)) \leq Pr(U-X|D^-(y)) . \end{aligned}$$

Condition (i) says that if objects x possesses properties related to attributes from AT at least as object y, i.e., $x \succsim y$, then the probability that x belongs to X has to be not smaller than the probability that y belongs to X. Analogously, Condition (ii) says that since $x \succsim y$, then the probability that x does not belong to X should not be greater than the probability that y does not belong to X. Observe that (ii) can be written also as

$$(ii)\ Pr(X|D^-(x)) \geq Pr(X|D^-(y)) .$$

These probabilities are unknown but can be estimated from data. For each $X \subseteq U$, we have a binary problem of estimating the conditional probabilities $Pr(X|D^+(x)) = 1 - Pr(U-X|D^+(x))$ and the conditional probabilities $Pr(X|D^-(x)) = 1 - Pr(U-X|D^-(x))$. It can be solved by *isotonic regression* [24.25]. For $X \subseteq U$ and for any $x \in U$, let $y(x, X) = 1$ if $x \in X$, otherwise $y(x, X) = 0$. Then one can choose estimates $Pr^*(X|D^+(x))$ and $Pr^*(X|D^-(x))$ with $Pr^*(X|D^+(x))$ and $Pr^*(X|D^-(x))$ which minimize the squared distance to the class assignment $y(x, X)$, subject to the monotonicity constraints related to the dominance relation $\succsim$ on the attributes from AT (see also Sect. 22.3.3 in Chap. 22)

$$\begin{aligned} &\text{Minimize} \\ &\sum_{x \in U} (y(x,X) - Pr(X|D^+(x)))^2 \\ &+ (y(x,X) - Pr(X|D^-(x)))^2 \\ &\text{subject to} \\ &Pr(X|D^+(x)) \geq Pr(X|D^+(z)) \text{ and} \\ &Pr(X|D^-(x)) \geq Pr(X|D^-(z)) \text{ if } x \succsim z, \\ &\quad \text{for all } x, z \in U . \end{aligned}$$

Then, stochastic α-lower approximations of $X \subseteq U$ can be defined as

$$\underline{P}^{\alpha}(X) = \{x \in U : Pr(X|D^{+}(x)) \geq \alpha\} ,$$
$$\underline{P}^{\alpha}(U-X) = \{x \in U : Pr(U-X|D^{-}(x)) \geq \alpha\} .$$

Replacing the unknown probabilities

$$Pr(X|D^{+}(x))$$

and

$$Pr(U-X|D^{-}(x))$$

by their estimates

$$Pr^{*}(X|D^{+}(x))$$

and

$$Pr^{*}(U-X|D^{-}(x))$$

obtained from isotonic regression, we get

$$\underline{P}^{\alpha}(X) = \left\{x \in U : Pr^{*}(X|D^{+}(x)) \geq \alpha\right\} ,$$
$$\underline{P}^{\alpha}(U-X) = \left\{x \in U : Pr^{*}(U-X|D^{-}(x)) \geq \alpha\right\} ,$$

where parameter $\alpha \in [0.5, 1]$ controls the allowed amount of inconsistency.

Solving isotonic regression requires $O(|U|^4)$ time, but a good heuristic needs only $O(|U|^2)$.

In fact, as shown in [24.25] and recalled in Sect. 22.3.3 in Chap. 22, we do not really need to know the probability estimates to obtain stochastic lower approximations. We only need to know for which object $x \in U$, $Pr^{*}(X|D^{+}(x)) \geq \alpha$ and for which $x \in U$, $Pr^{*}(U-X|D^{-}(x)) \geq \alpha$ (i.e., $Pr^{*}(X|D^{-}(x)) \leq 1-\alpha$). This can be found by solving a linear programming (reassignment) problem.

As before, $y(x,X) = 1$ if $x \in X$, otherwise $y(x,X) = 0$. Let $d(x,X)$ be the decision variable which determines a new class assignment for object x. Then, reassign objects to X if $d^{*}(x,X) = 1$, and to $U-X$ if $d^{*}(x,X) = 0$, such that the new class assignments are consistent with the dominance principle, where $d^{*}(x,X)$ results from solving the following linear programming problem

$$\begin{aligned}&\text{Minimize} \sum_{x \in U} w_{y(x,X)} |y(x,X) - d(x,X)| \\ &\text{subject to } d(x,X) \geq d(z,X) \text{ if } x \succsim z \\ &\qquad\qquad \text{for all } x, z \in U\end{aligned}$$

where w_1 and w_0 are arbitrary positive weights.

Due to unimodularity of the constraint matrix, the optimal solution of this linear programming problem is always integer, i. e., $d^{*}(x,X) \in \{0,1\}$. For all objects consistent with the dominance principle, $d^{*}(x,X) = y(x,X)$. If we set $w_0 = \alpha$ and $w_1 = \alpha - 1$, then the optimal solution $d^{*}(x,X)$ satisfies: $d^{*}(x,X) = 1 \Leftrightarrow Pr^{*}(X|D^{+}(x)) \geq \alpha$. If we set $w_0 = 1-\alpha$ and $w_1 = \alpha$, then the optimal solution $d^{*}(x,X)$ satisfies: $d^{*}(x,X) = 0 \Leftrightarrow Pr^{*}(X|D^{-}(x)) \leq 1-\alpha$.

Solving the reassignment problem twice, we can obtain the lower approximations $\underline{P}^{\alpha}(X)$, $\underline{P}^{\alpha}(U-X)$, without knowing the probability estimates.

24.9.3 Three-Way Decisions: Interpreting the Three Regions in the Case of Dominance-Based Rough Sets

In this section, we present an interpretation of dominance-based rough set three regions taking into consideration the framework of three-way decisions.

In an information table, with respect to a subset of attributes $A \subseteq AT$, an object x induces logic formulae

$$\bigwedge_{a \in A} I_a(x) \succsim_a v_a , \tag{24.84}$$

$$\bigwedge_{a \in A} v_a \succsim_a I_a(x) , \tag{24.85}$$

where $I_a(x), v_a \in V_a$ and

- The atomic formula $v_a \succsim_a I_a(x)$ indicates that object x taking value $I_a(x)$ on attribute a possess a property related to a not more than any object y taking value $I_a(y) = v_a$ on attribute a.
- The atomic formula $I_a(x) \succsim_a v_a$ indicates that object x taking value $I_a(x)$ on attribute a possess a property related to a not less than any object y taking value $I_a(y) = v_a$ on attribute a.

Thus, an object y satisfies the formula

$$\bigwedge_{a \in A} I_a(x) \succsim_a v_a \text{ , if } I_a(y) \succsim v_a \text{ for all } a \in A ,$$

that is,

$$\left(y \models \bigwedge_{a \in A} I_a(x) \succsim_a v_a \Longleftrightarrow \forall a \in A,\ \ (I_a(y) \succsim_a v_a\right) . \tag{24.86}$$

Analogously, an object y satisfies the formula

$$\bigwedge_{a\in A} v_a \succsim_a I_a(x)\text{ , if } v_a \succsim I_a(y) \quad \text{for all } a \in A\text{ ,}$$

that is,

$$\left(y \models \bigwedge_{a\in A} v_a \succsim_a I_a(x) \Longleftrightarrow \forall a \in A,\quad (v_a \succsim_a I_a(y)\right)\text{ .} \tag{24.87}$$

With these notations, we are ready to interpret upward and downward dominance-based rough set three regions.

From the upward and downward three regions, we can construct three classes of rules for classifying an object, called the upward and downward positive, negative, and boundary rules.

They are expressed in the following forms: for $y \in U$,

- Positive rule induced by an upward cone

$$D^+(x) \subseteq \text{POS}^+_{(\alpha,\beta)}(X):$$
$$\text{if } y \models \bigwedge_{a\in A} I_a(x) \succsim_a v_a\text{ , accept } y \in X\text{ ,}$$

- Negative rule induced by the complement of an upward cone

$$U - D^+(x) \subseteq \text{NEG}^+_{(\alpha,\beta)}(X):$$
$$\text{if } y \models \neg \bigwedge_{a\in A} I_a(x) \succsim_a v_a\text{ , reject } y \in X\text{ ,}$$

- Boundary rule induced by an upward cone $D^+(x)$ and its complement $U - D^+(x)$ such that

$$D^+(x) \not\subseteq \text{POS}^+_{(\alpha,\beta)}(X)$$
$$\text{and } (U - D^+(x)) \not\subseteq \text{NEG}^+_{(\alpha,\beta)}(X):$$
$$\text{if } y \models \bigwedge_{a\in A} I_a(x) \succsim_a v_a \wedge \neg \bigwedge_{a\in A} I_a(x) \succsim_a u_a,$$
$$\text{neither accept nor reject } y \in X\text{ ,}$$

- Positive rule induced by an downward cone

$$D^-(x) \subseteq \text{POS}^-_{(\alpha,\beta)}(X):$$
$$\text{if } y \models \bigwedge_{a\in A} v_a \succsim_a I_a(x)\text{, accept } y \in X\text{ ,}$$

- Negative rule induced by the complement of a downward cone

$$U - D^-(x) \subseteq \text{NEG}^-_{(\alpha,\beta)}(X):$$
$$\text{if } y \models \neg \bigwedge_{a\in A} v_a \succsim_a I_a(x)\text{, reject } y \in X\text{ ,}$$

- Boundary rule induced by a downward cone $D^-(x)$ and its complement $U - D^-(x)$ such that

$$D^-(x) \not\subseteq \text{POS}^-_{(\alpha,\beta)}(X)$$
$$\text{and } U - D^-(x) \not\subseteq \text{NEG}^-_{(\alpha,\beta)}(X)\ :$$
$$\text{if } y \models \bigwedge_{a\in A} v_a \succsim_a I_a(x) \wedge \neg \bigwedge_{a\in A} u_a \succsim_a I_a(x),$$
$$\text{neither accept nor reject } y \in X\text{ .}$$

The three types of rules have a semantic interpretations analogous to those induced by probabilistic rough sets based on indiscernibility presented in Sect. 24.5.3. Let us consider the rules related to POS^+ and NEG^+. A positive rule allows us to *accept* an object y to be a member of X, because y has a higher probability of being in X due to the facts that $y \in D^+(x)$ and $Pr(X|D^+(x)) \geq \alpha^+$. A negative rule enables us to *reject* an object y to be a member of X, because y has lower probability of being in X due to the facts that $y \in D^+(x)$ and $Pr(X|D^+(x)) \leq \beta^+$. When the probability of y being in X is neither high nor low, a boundary rule makes a noncommitment decision.

The error rate of a positive rule is given by $1 - Pr(X|D^+(x))$, which, by definition of the three regions, is at or below $1 - \alpha^+$. The error rate of negative rule is given by $Pr(X|D^+(x))$ and is at or below β^+. The cost of a positive rule is $\lambda^+_{PP} Pr(X|D^+(x)) + \lambda^+_{PN}(1 - Pr(X|D^+(x)))$ and is bounded above by $\alpha^+\lambda^+_{PP} + (1 - \alpha^+)\lambda^+_{PN}$. The cost of a negative rule is $\lambda^+_{NP} Pr(X|D^+(x)) + \lambda^+_{NN}(1 - Pr(X|D^+(x)))$ and is bounded above by $\beta^+\lambda^+_{NP} + (1 - \beta^+)\lambda^+_{NN}$.

24.10 Conclusions

A basic probabilistic rough set model is formulated by using a pair of thresholds on conditional probabilities, which leads to flexibility and robustness when performing classification or decision-making tasks. Three theories are the supporting pillars of probabilistic rough sets. Bayesian decision theory enables us to determine and interpret the required thresholds by using more operable notions such as loss, cost, risk, etc. Bayesian inference ensures us to estimate the conditional probability accurately. A theory of three-way decisions allows us to make a wise decision in the presence of incomplete or insufficient information.

Other probabilistic rough set models have also been described. We have shown how a probabilistic approach can be applied when information related to some order. The order concerns degrees in which an object has some properties related to considered attributes. This kind of order can be handled by the well-known rough set extension called dominance-based rough set approach.

One may expect a continuous growth of interest in probabilistic approaches to rough sets. An important task is to examine fully, in the light of three fundamental issues concerning the basic model, the semantics of each model, in order to identify its limitations and appropriate areas of applications.

References

24.1 Z. Pawlak: Rough set, Int. J. Inf. Comput. Sci. **11**, 341–356 (1982)
24.2 Z. Pawlak: *Rough Sets: Theoretical Aspects of Reasoning About Data* (Kluwer, Dordrecht 1991)
24.3 W. Marek, Z. Pawlak: Information storage and retrieval systems: mathematical foundations, Theor. Comput. Sci. **1**, 331–354 (1976)
24.4 Y.Y. Yao: A note on definability and approximations. In: *Transactions on Rough Sets VII*, Lecture Notes in Computer Science, Vol. 4400, ed. by J.F. Peters, A. Skowron, V.W. Marek, E. Orlowska, R. Słowiński, W. Ziarko (Springer, Heidelberg 2007) pp. 274–282
24.5 Y.Y. Yao: Probabilistic approaches to rough sets, Expert Syst. **20**, 287–297 (2003)
24.6 Y.Y. Yao: Probabilistic rough set approximations, Int. J. Approx. Reason. **49**, 255–271 (2008)
24.7 Z. Pawlak, S.K.M. Wong, W. Ziarko: Rough sets: Probabilistic versus deterministic approach, Int. J. Man-Mach. Stud. **29**, 81–95 (1988)
24.8 S. K. M. Wong, W. Ziarko: A probabilistic model of approximate classification and decision rules with uncertainty in inductive learning, Technical Report CS-85-23 (Department of Computer Science, University of Regina 1985)
24.9 S.K.M. Wong, W. Ziarko: INFER – an adaptive decision support system based on the probabilistic approximate classifications, Proc. 6th Int. Workshop on Expert Syst. Their Appl., Vol. 1 (1986) pp. 713–726
24.10 S.K.M. Wong, W. Ziarko: Comparison of the probabilistic approximate classification and the fuzzy set model, Fuzzy Sets Syst. **21**(3), 357–362 (1987)
24.11 Y.Y. Yao, S.K.M. Wong: A decision theoretic framework for approximating concepts, Int. J. Man-Mach. Stud. **37**, 793–809 (1992)
24.12 Y.Y. Yao, S.K.M. Wong, P. Lingras: A decision-theoretic rough set model. In: *Methodologies for Intelligent Systems*, Vol. 5, ed. by Z.W. Ras, M. Zemankova, M.L. Emrich (North-Holland, New York 1990) pp. 17–24
24.13 J.D. Katzberg, W. Ziarko: Variable precision rough sets with asymmetric bounds. In: *Rough Sets, Fuzzy Sets and Knowledge Discovery*, ed. by W. Ziarko (Springer, Heidelberg 1994) pp. 167–177
24.14 W. Ziarko: Variable precision rough set model, J. Comput. Syst. Sci. **46**, 39–59 (1993)
24.15 D. Ślęzak, W. Ziarko: Bayesian rough set model, Proc. Found. Data Min. (FDM 2002) (2002) pp. 131–135
24.16 D. Ślęzak, W. Ziarko: Variable precision Bayesian rough set model, Rough Sets, Fuzzy Sets, Data Minging and Granular Comput. (RSFGrC 2013), Lect. Notes Comput. Sci. (Lect. Notes Artif. Intel.), Vol. 2639, ed. by G.Y. Wang, Q. Liu, Y.Y. Yao, A. Skowron (Springer, Heidelberg 2003) pp. 312–315
24.17 D. Ślęzak, W. Ziarko: The investigation of the Bayesian rough set model, Int. J. Approx. Reason. **40**, 81–91 (2005)
24.18 H.Y. Zhang, J. Zhou, D.Q. Miao, C. Gao: Bayesian rough set model: a further investigation, Int. J. Approx. Reason. **53**, 541–557 (2012)
24.19 S. Greco, B. Matarazzo, R. Słowiński: Rough membership and Bayesian confirmation measures for parameterized rough sets. In: *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing*, Lecture Notes in Computer Science, Vol. 3641, ed. by D. Ślęzak, G.Y. Wang, M. Szczuka, I. Duntsch, Y.Y. Yao (Springer, Heidelberg 2005) pp. 314–324
24.20 S. Greco, B. Matarazzo, R. Słowiński: Parameterized rough set model using rough membership and Bayesian confirmation measures, Int. J. Approx. Reason. **49**, 285–300 (2008)
24.21 N. Azam, J.T. Yao: Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets, Int. J. Approx. Reason. **55**, 142–155 (2014)

24.22 J.P. Herbert, J.T. Yao: Game-theoretic rough sets, Fundam. Inf. **108**, 267–286 (2011)
24.23 S. Greco, B. Matarazzo, R. Słowiński, J. Stefanowski: Variable consistency model of dominance-based rough set approach. In: *Rough Sets and Current Trends in Computing*, Lecture Notes in Computer Science, Vol. 2005, ed. by W. Ziarko, Y.Y. Yao (Springer, Heidelberg 2001) pp. 170–181
24.24 J. Błaszczyński, S. Greco, R. Słowiński, M. Szelag: Monotonic variable consistency rough set approaches, Int. J. Approx. Reason. **50**, 979–999 (2009)
24.25 W. Kotłowski, K. Dembczyński, S. Greco, R. Słowiński: Stochastic dominance-based rough set model for ordinal classification, Inf. Sci. **178**, 4019–4037 (2008)
24.26 B. Zhou, Y.Y. Yao: Feature selection based on confirmation-theoretic rough sets. In: *Rough Sets and Current Trends in Computing*, Lecture Notes in Computer Science, Vol. 8536, ed. by C. Cornelis, M. Kryszkiewicz, D. Ślęzak, E.M. Ruiz, R. Bello, L. Shang (Springer, Heidelberg 2014) pp. 181–188
24.27 X.F. Deng, Y.Y. Yao: An information-theoretic interpretation of thresholds in probabilistic rough sets. In: *Rough Sets and Knowledge Technology*, Lecture Notes in Computer Science, Vol. 7414, ed. by T.R. Li, H.S. Nguyen, G.Y. Wang, J. Grzymala-Busse, R. Janicki (Springer, Heidelberg 2012) pp. 369–378
24.28 B. Zhou, Y.Y. Yao: Comparison of two models of probabilistic rough sets. In: *Rough Sets and Knowledge Technology*, Lecture Notes in Computer Science, Vol. 8171, ed. by P. Lingras, M. Wolski, C. Cornelis, S. Mitra, P. Wasilewski (Springer, Heidelberg 2013) pp. 121–132
24.29 J.W. Grzymala-Busse: Generalized parameterized approximations. In: *Rough Sets and Knowledge Technology*, Lecture Notes in Computer Science, Vol. 6954, ed. by J.T. Yao, S. Ramanna, G.Y. Wang, Z. Suraj (Springer, Heidelberg 2011) pp. 36–145
24.30 J.W. Grzymala-Busse: Generalized probabilistic approximations. In: *Transactions on Rough Sets*, Lecture Notes in Computer Science, Vol. 7736, ed. by J.F. Peters, A. Skowron, S. Ramanna, Z. Suraj, X. Wang (Springer, Heidelberg 2013) pp. 1–16
24.31 S. Greco, B. Matarazzo, R. Słowiński: Rough sets theory for multicriteria decision analysis, Eur. J. Oper. Res. **129**, 1–47 (2001)
24.32 S. Greco, B. Matarazzo, R. Słowiński: Decision rule approach. In: *Multiple Criteria Decision Analysis: State of the Art Surveys*, ed. by J.R. Figueira, S. Greco, M. Ehrgott (Springer, Berlin 2005) pp. 507–562
24.33 R. Słowiński, S. Greco, B. Matarazzo: Rough sets in decision making. In: *Encyclopedia of Complexity and Systems Science*, ed. by R.A. Meyers (Springer, New York 2009) pp. 7753–7786
24.34 R. Słowiński, S. Greco, B. Matarazzo: Rough set and rule-based multicriteria decision aiding, Pesqui. Oper. **32**, 213–269 (2012)
24.35 Y.Y. Yao: Relational interpretations of neighborhood operators and rough set approximation operators, Inf. Sci. **111**, 239–259 (1998)
24.36 Y.Y. Yao: Information granulation and rough set approximation, Int. J. Intell. Syst. **16**, 87–104 (2001)
24.37 Y.Y. Yao, Y.H. Chen: Subsystem based generalizations of rough set approximations. In: *Foundations of Intelligent Systems*, Lecture Notes in Computer Science, Vol. 3488, ed. by M.S. Hacid, N.V. Murray, Z.W. Raś, S. Tsumoto (Springer, Heidelberg 2005) pp. 210–218
24.38 Y.Y. Yao, X.F. Deng: Quantitative rough sets based on subsethood measures, Inf. Sci. **267**, 702–715 (2014)
24.39 H.X. Li, X.Z. Zhou, T.R. Li, G.Y. Wang, D.Q. Miao, Y.Y. Yao: *Decision-Theoretic Rough Set Theory and Recent Progress* (Science Press, Beijing 2011)
24.40 H. Yu, G.Z. Liu, Y.G. Wang: An automatic method to determine the number of clusters using decision-theoretic rough set, Int. J. Approx. Reason. **55**, 101–115 (2014)
24.41 F. Li, M. Ye, D.X. Chen: An extension to rough c-means clustering based on decision-theoretic rough sets model, Int. J. Approx. Reason. **55**, 116–129 (2014)
24.42 J. Li, T.P.X. Yang: An axiomatic characterization of probabilistic rough sets, Int. J. Approx. Reason. **55**, 130–141 (2014)
24.43 X.Y. Jia, Z.M. Tang, W.H. Liao, L. Shang: On an optimization representation of decision-theoretic rough set model, Int. J. Approx. Reason. **55**, 156–166 (2014)
24.44 F. Min, Q.H. Hu, W. Zhu: Feature selection with test cost constraint, Int. J. Approx. Reason. **55**, 167–179 (2014)
24.45 J.W. Grzymala-Busse, G.P. Clark, M. Kuehnhausen: Generalized probabilistic approximations of incomplete data, Int. J. Approx. Reason. **55**, 180–196 (2014)
24.46 D. Liu, T.R. Li, D.C. Liang: Incorporating logistic regression to decision-theoretic rough sets for classifications, Int. J. Approx. Reason. **55**, 197–210 (2014)
24.47 B. Zhou: Multi-class decision-theoretic rough sets, Int. J. Approx. Reason. **55**, 211–224 (2014)
24.48 H.Y. Qian, H. Zhang, L.Y. Sang, Y.J. Liang: Multigranulation decision-theoretic rough sets, Int. J. Approx. Reason. **55**, 225–237 (2014)
24.49 P. Lingras, M. Chen, Q.D. Miao: Qualitative and quantitative combinations of crisp and rough clustering schemes using dominance relations, Int. J. Approx. Reason. **55**, 238–258 (2014)
24.50 W.M. Shao, Y. Leung, Z.W. Wu: Rule acquisition and complexity reduction in formal decision contexts, Int. J. Approx. Reason. **55**, 259–274 (2014)
24.51 J.T. Yao, X.X. Li, G. Peters: Decision-theoretic rough sets and beyond, Int. J. Approx. Reason. **55**, 9–100 (2014)
24.52 X.Y. Zhang, D.Q. Miao: Two basic double-quantitative rough set models of precision and grade and their investigation using granular

computing, Int. J. Approx. Reason. **54**, 1130–1148 (2013)
24.53 W. Ziarko: Probabilistic approach to rough sets, Int. J. Approx. Reason. **49**, 272–284 (2008)
24.54 B. Fitelson: Studies in Bayesian Confirmation Theory, Ph.D. Thesis (University of Wisconsin, Madison 2001)
24.55 R. Festa: Bayesian confirmation. In: *Experience, Reality, and Scientific Explanation*, ed. by M. Galavotti, A. Pagnini (Kluwer, Dordrecht 1999) pp. 55–87
24.56 S. Greco, Z. Pawlak, R. Słowiński: Can Bayesian confirmation measures be useful for rough set decision rules?, Eng. Appl. Artif. Intell. **17**, 345–361 (2004)
24.57 S. Greco, R. Słowiński, I. Szczęch: Properties of rule interestingness measures and alternative approaches to normalization of measures, Inf. Sci. **216**, 1–16 (2012)
24.58 Y.Y. Yao: Two semantic issues in a probabilistic rough set model, Fundam. Inf. **108**, 249–265 (2011)
24.59 Y.Y. Yao, B. Zhou: Naive Bayesian rough sets. In: *Rough Sets and Knowledge Technology*, Lecture Notes in Computer Science, Vol. 6401, ed. by J. Yu, S. Greco, P. Lingras, G.Y. Wang, A. Skowron (Springer, Heidelberg 2010) pp. 719–726
24.60 D.C. Liang, D. Liu, W. Pedrycz, P. Hu: Triangular fuzzy decision-theoretic rough sets, Int. J. Approx. Reason. **54**, 1087–1106 (2013)
24.61 H.X. Li, X.Z. Zhou: Risk decision making based on decision-theoretic rough set: a three-way view decision model, Int. J. Comput. Intell. Syst. **4**, 1–11 (2011)
24.62 D. Liu, T.R. Li, D. Ruan: Probabilistic model criteria with decision-theoretic rough sets, Inf. Sci. **181**, 3709–3722 (2011)
24.63 K. Dembczyński, S. Greco, W. Kotłowski, R. Słowiński: Statistical model for rough set approach to multicriteria classification. In: *Knowledge Discoveery in Databases*, Lecture Notes in Computer Science, Vol. 4702, ed. by J.N. Kok, J. Koronacki, R. de Lopez Mantaras, S. Matwin, D. Mladenic, A. Skowron (Springer, Heidelberg 2007) pp. 164–175
24.64 Y.Y. Yao: An outline of a theory of three-way decisions. In: *Rough Sets and Current Trends in Computing*, Lecture Notes in Computer Science, Vol. 7413, ed. by J.T. Yao, Y. Yang, R. Słowiński, S. Greco, H.X. Li, S. Mitra, L. Polkowski (Springer, Heidelberg 2012) pp. 1–17
24.65 Y.Y. Yao: Three-way decision: an interpretation of rules in rough set theory. In: *Rough Sets and Knowledge Technology*, Lecture Notes in Computer Science, Vol. 5589, ed. by P. Wen, Y.F. Li, L. Polkowski, Y.Y. Yao, S. Tsumoto, G.Y. Wang (Springer, Heidelberg 2009) pp. 642–649
24.66 Y.Y. Yao: Three-way decisions with probabilistic rough sets, Inf. Sci. **180**, 341–353 (2010)
24.67 Y.Y. Yao: The superiority of three-way decisions in probabilistic rough set models, Inf. Sci. **181**, 1080–1096 (2011)
24.68 X.Y. Jia, L. Shang, X.Z. Zhou, J.Y. Liang, D.Q. Miao, G.Y. Wang, T.R. Li, Y.P. Zhang: *Theory of Three-Way Decisions and Application* (Nanjing Univ. Press, Nanjing 2012)
24.69 D. Liu, T.R. Li, D.Q. Miao, G.Y. Wang, J.Y. Liang: *Three-Way Decisions and Granular Computing* (Science Press, Beijing 2013)
24.70 J.R. Figueira, S. Greco, M. Ehrgott: *Multiple Criteria Decision Analysis: State of the Art Surveys* (Springer, Berlin 2005)
24.71 S. Greco, B. Matarazzo, R. Słowiński: A new rough set approach to evaluation of bankruptcy risk. In: *Rough Fuzzy and Fuzzy Rough Sets*, ed. by C. Zopounidis (Kluwer, Dordrecht 1998) pp. 121–136
24.72 S. Greco, B. Matarazzo, R. Słowiński: The use of rough sets and fuzzy sets in MCDM. In: *Multicriteria Decision Making*, Int. Ser. Opear. Res. Manage. Sci., Vol. 21, ed. by T. Gal, T. Stewart, T. Hanne (Kluwer, Dordrecht 1999) pp. 397–455
24.73 S. Greco, B. Matarazzo, R. Słowiński: Extension of the rough set approach to multicriteria decision support, INFOR **38**, 161–196 (2000)
24.74 S. Greco, B. Matarazzo, R. Słowiński: Rough sets methodology for sorting problems in presence of multiple attributes and criteria, Eur. J. Oper. Res. **138**, 247–259 (2002)
24.75 S. Greco, R. Słowiński, Y. Yao: Bayesian decision theory for dominance-based rough set approach. In: *Rough Sets and Knowledge Technology*, Lecture Notes in Computer Science, Vol. 4481, ed. by J.T. Yao, P. Lingras, W.Z. Wu, M. Szczuka, N. Cercone (Springer, Heidelberg 2007) pp. 134–141