ROUGH SETS AND DATA MINING

Analysis for Imprecise Data

T. Y. LIN
San Jose State University
San Jose, California, USA

N. CERCONE
University of Regina
Regina, Saskatchewan, CANADA

KLUWER ACADEMIC PUBLISHERS
Boston/London/Dordrecht

1997

3

A REVIEW OF ROUGH SET MODELS

Y.Y. Yao*, S.K.M. Wong**, and T.Y. Lin***

** Department of Computer Science, Lakehead University
Thunder Bay, Ontario, Canada P7B 5E1*

*** Department of Computer Science, University of Regina
Regina, Saskatchewan, Canada S4S 0A2*

**** Department of Mathematics and Computer Science
San Jose State University, San Jose, California 95192*

ABSTRACT

Since introduction of the theory of rough set in early eighties, considerable work has been done on the development and application of this new theory. The paper provides a review of the Pawlak rough set model and its extensions, with emphasis on the formulation, characterization, and interpretation of various rough set models.

1 INTRODUCTION

In early eighties, Pawlak [22] introduced the theory of rough sets as an extension of set theory for the study of intelligent systems characterized by insufficient and incomplete information [22, 23, 26]. It is motivated by the practical needs in classification and concept formation [27]. One may regard the theory of rough sets to be complementary to other generalizations of set theory, such as fuzzy sets and multisets [6, 24, 27, 42]. In recent years, there has been a fast growing interest in this new emerging theory. The successful applications of the rough set model in a variety of problems have amply demonstrated its usefulness and versatility [13, 15, 25, 33, 50].

The main objective of this paper is to present a review of the standard rough set model and its extensions, and to give some new results. Our emphasis will be on the formulation, characterization, and interpretation of various rough set models. We group existing rough set models into two major classes, the algebraic and probabilistic rough set models, depending on whether statistical information is used. In the algebraic class, we examine different rough set models in relation to modal logic, graded rough set models, rough set models over two universes, and rough set models over Boolean algebras. In the probabilistic class, we analyze rough membership functions and variable precision rough set models. More importantly, the probabilistic rough set models are justified based on the framework of decision theory.

In this paper, binary relations are used as a primitive notion. Rough set models are built and investigated based on various binary relations. Our aim is not to provide a complete and exhaustive summary of all works on rough set models. We only review existing works that fall in the framework we intent to establish based on binary relations. Many important studies, such as the construction of rough set model based on a covering of the universe [48] and algebraic study of rough set models [30, 37], are not covered in this paper.

2 ALGEBRAIC ROUGH SET MODELS

This section reviews the Pawlak rough set model and presents its extensions and interpretations.

2.1 Pawlak rough set model

Let U denote a finite and non-empty set called the universe, and let $\Re \subseteq U \times U$ denote an equivalence relation on U. The pair $apr = (U, \Re)$ is called an approximation space. The equivalence relation $\Re$ partitions the set U into disjoint subsets. Such a partition of the universe is denoted by $U/\Re$. If two elements x, y in U belong to the same equivalence class, we say that x and y are indistinguishable. The equivalence classes of $\Re$ and the empty set $\emptyset$ are called the elementary or atomic sets in the approximation space $apr = (U, \Re)$. The union of one or more elementary sets is called a composed set. The family of all composed sets, including the empty set, is denoted by $\mathrm{Com}(apr)$, which forms a Boolean algebra.

The equivalence relation and the induced equivalence classes may be regarded as the available information or knowledge about the objects under consideration. Given an arbitrary set $X \subseteq U$, it may be impossible to describe X precisely using the equivalence classes of $\Re$. That is, the available information is not sufficient to give a precise representation of X. In this case, one may characterize X by a pair of lower and upper approximations:

$$\begin{aligned} \underline{apr}(X) &= \bigcup_{[x]_\Re \subseteq X} [x]_\Re, \\ \overline{apr}(X) &= \bigcup_{[x]_\Re \cap X \neq \emptyset} [x]_\Re, \end{aligned} \tag{1.1}$$

where

$$[x]_\Re = \{y \mid x \Re y\}, \tag{1.2}$$

is the equivalence class containing x. The lower approximation $\underline{apr}(X)$ is the union of all the elementary sets which are subsets of X. It is the largest composed set contained in X. The upper approximation $\overline{apr}(X)$ is the union of all the elementary sets which have a non-empty intersection with X. It is the smallest composed set containing X. An element in the lower approximation necessarily belongs to X, while an element in the upper approximation possibly belongs to X. We can also express lower and upper approximations as follow:

$$\begin{aligned} \underline{apr}(X) &= \{x \mid [x]_\Re \subseteq X\} \\ \overline{apr}(X) &= \{x \mid [x]_\Re \cap X \neq \emptyset\}. \end{aligned} \tag{1.3}$$

That is, an element of U necessarily belongs to X if all its equivalent elements belong to X; it is possibly belongs to X if at least one of its equivalent elements belongs to X.

For any subsets $X, Y \subseteq U$, the lower approximation $\underline{apr}$ satisfies properties:

$$\begin{array}{ll}
\text{(AL1)} & \underline{apr}(X) = \sim\overline{apr}(\sim X), \\
\text{(AL2)} & \underline{apr}(U) = U, \\
\text{(AL3)} & \underline{apr}(X \cap Y) = \underline{apr}(X) \cap \underline{apr}(Y), \\
\text{(AL4)} & \underline{apr}(X \cup Y) \supseteq \underline{apr}(X) \cup \underline{apr}(Y), \\
\text{(AL5)} & X \subseteq Y \Longrightarrow \underline{apr}(X) \subseteq \underline{apr}(Y), \\
\text{(AL6)} & \underline{apr}(\emptyset) = \emptyset, \\
\text{(AL7)} & \underline{apr}(X) \subseteq X, \\
\text{(AL8)} & X \subseteq \underline{apr}(\overline{apr}(X)), \\
\text{(AL9)} & \underline{apr}(X) \subseteq \underline{apr}(\underline{apr}(X)), \\
\text{(AL10)} & \overline{apr}(X) \subseteq \underline{apr}(\overline{apr}(X)),
\end{array}$$

and the upper approximation $\overline{apr}$ satisfies properties:

$$\text{(AU1)} \quad \overline{apr}(X) = \sim\underline{apr}(\sim X),$$

$$
\begin{array}{ll}
\text{(AU2)} & \overline{apr}(\emptyset) = \emptyset, \\
\text{(AU3)} & \overline{apr}(X \cup Y) = \overline{apr}(X) \cup \overline{apr}(Y), \\
\text{(AU4)} & \overline{apr}(X \cap Y) \subseteq \overline{apr}(X) \cap \overline{apr}(Y), \\
\text{(AU5)} & X \subseteq Y \Longrightarrow \overline{apr}(X) \subseteq \overline{apr}(Y), \\
\text{(AU6)} & \overline{apr}(U) = U, \\
\text{(AU7)} & X \subseteq \overline{apr}(X), \\
\text{(AU8)} & \overline{apr}(\underline{apr}(X)) \subseteq X, \\
\text{(AU9)} & \overline{apr}(\overline{apr}(X)) \subseteq \overline{apr}(X), \\
\text{(AU10)} & \overline{apr}(\underline{apr}(X)) \subseteq \underline{apr}(X),
\end{array}
$$

where $\sim X = U - X$ denotes the set complement of X. The lower and upper approximations may be viewed as two operators on the universe [14]. Properties (AL1) and (AU1) state that two approximation operators are dual operators. Hence, properties with the same number may be regarded as dual properties. These properties are not independent.

Based on the lower and upper approximations of a set $X \subseteq U$, the universe U can be divided into three disjoint regions, the positive region $\mathrm{POS}(X)$, the negative region $\mathrm{NEG}(X)$, and the boundary region $\mathrm{BND}(X)$:

$$
\begin{aligned}
\mathrm{POS}(X) &= \underline{apr}(X), \\
\mathrm{NEG}(X) &= U - \overline{apr}(X), \\
\mathrm{BND}(X) &= \overline{apr}(X) - \underline{apr}(X).
\end{aligned} \tag{1.4}
$$

Figure 1 illustrates the approximation of a set X, and the positive, negative and boundary regions. Each small rectangle represent an equivalence class. From this figure, we have the following observations. One can say with certainty that any element $x \in \mathrm{POS}(X)$ belongs to X, and that any element $x \in \mathrm{NEG}(X)$ does not belong to X. The upper approximation of a set X is the union of the positive and boundary regions, namely, $\overline{apr}(X) = \mathrm{POS}(X) \cup \mathrm{BND}(X)$. One cannot decide with certainty whether or not an element $x \in \mathrm{BND}(X)$ belongs to X. For arbitrary element $x \in \overline{apr}(X)$, one can only conclude that x possibly belongs to X.

An important concept related to lower and upper approximations is the accuracy of the approximation of a set [22]. Yao and Lin [44] have shown that the accuracy of approximation can be interpreted using the well-known Marczewski-Steinhaus metric, or MZ metric for short. For two sets X and Y, the Marczewski-Steinhaus metric measures the distance between two sets [18]:

$$
\begin{aligned}
D(X,Y) &= \frac{|X \Delta Y|}{|X \cup Y|} \\
&= 1 - \frac{|X \cap Y|}{|X \cup Y|},
\end{aligned} \tag{1.5}
$$

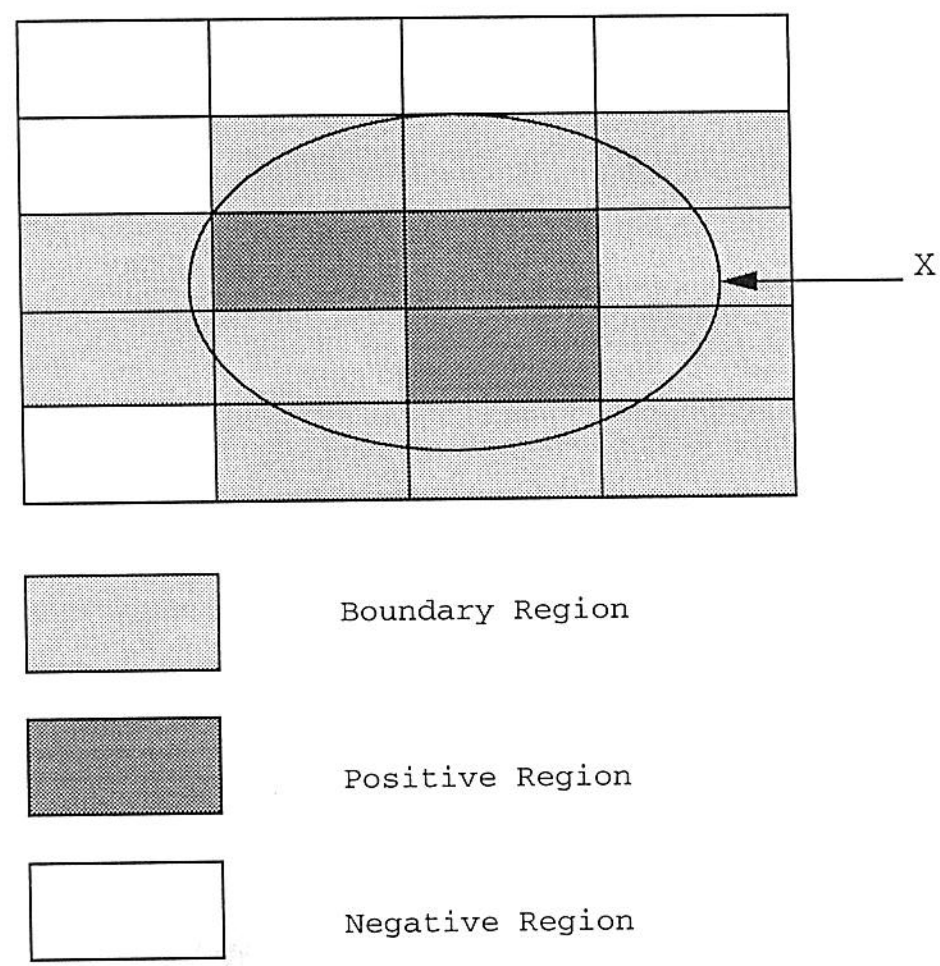

Figure 1 Positive, boundary and negative regions of a set X

where $X \Delta Y = (X \cup Y) - (X \cap Y)$ denotes the symmetric difference between two sets X and Y, and $|\cdot|$ the cardinality of a set. It reaches the maximum value of 1 if X and Y are disjoint, i.e., they are totally different, and it reaches the minimum value of 0 if X and Y are exactly the same. The quantity,

$$S(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}, \tag{1.6}$$

may be interpreted as a measure of similarity or closeness between X and Y. By applying the MZ metric to the lower and upper approximations, we have:

$$\begin{aligned} D(\underline{apr}(X), \overline{apr}(X)) &= 1 - \frac{|\underline{apr}(X) \cap \overline{apr}(X)|}{|\underline{apr}(X) \cup \overline{apr}(X)|} \\ &= 1 - \frac{|\underline{apr}(X)|}{|\overline{apr}(X)|}, \end{aligned} \tag{1.7}$$

The distance function defined above is indeed the *inverse* function of the accuracy of rough set approximation proposed by Pawlak [22], namely,

$$\begin{aligned} \rho(X) &= 1 - D(\underline{apr}(X), \overline{apr}(X)) \\ &= \frac{|\underline{apr}(X)|}{|\overline{apr}(X)|} \\ &= S(\underline{apr}(X), \overline{apr}(X)). \end{aligned} \tag{1.8}$$

For the empty set $\emptyset$, we define $\rho(\emptyset) = 1$. If X is a composed set, then $\rho(X) = 1$. If X is not composed set, then $0 \leq \rho(X) < 1$.

In the Pawlak rough set model, an arbitrary set is described by a pair of lower and upper approximations. Several different interpretations of the concepts of rough sets have been proposed. The interpretation suggested by Iwinski [11] views a rough set as a pair of composed sets, and the original proposal of Pawlak regards a rough set as a family of sets having the same lower and/or upper approximation. Rough sets may also be described by using the notion of rough membership functions, which will be discussed in Section 3.

Given two composed sets $X_1, X_2 \in \mathrm{Com}(apr)$ with $X_1 \subseteq X_2$, Iwinski called the pair (X_1, X_2) an rough set [11]. In order to distinguish it from other definition, we call the pair an I-rough set. Let $R(apr)$ be the set of all I-rough sets. Set-theoretic operators on $R(apr)$ can be defined component-wise using standard set operators. For a pair of I-rough sets, we have:

$$\begin{aligned} (X_1, X_2) \cap (Y_1, Y_2) &= (X_1 \cap Y_1, X_2 \cap Y_2), \\ (X_1, X_2) \cup (Y_1, Y_2) &= (X_1 \cup Y_1, X_2 \cup Y_2). \end{aligned} \tag{1.9}$$

The intersection and union of two composed sets are still composed sets. The above operators are well defined, as the results are also I-rough sets. The system $(R(apr), \cap, \cup)$ is complete distributive lattice [11], with zero element $(\emptyset, \emptyset)$ and unit element (U, U). The associated order relation can be interpreted as I-rough set inclusion, which is defined by:

$$(X_1, X_2) \subseteq (Y_1, Y_2) \iff X_1 \subseteq Y_1 \text{ and } X_2 \subseteq Y_2. \tag{1.10}$$

The difference of I-rough sets can be defined as

$$(X_1, X_2) - (Y_1, Y_2) = (X_1 - Y_2, X_2 - Y_1), \tag{1.11}$$

which is an I-rough set. Finally, the I-rough set complement is given as:

$$\sim (X_1, X_2) = (U, U) - (X_1, X_2) = (\sim X_2, \sim X_1). \tag{1.12}$$

The complement is neither a Boolean complement nor a pseudocomplement in the lattice $(R(apr), \cap, \cup)$. The system $(R(apr), \cap, \cup, \sim, (\emptyset, \emptyset), (U, U))$ is called an I-rough set algebra.

In Pawlak's seminal paper, another interpretation of rough sets was introduced. Using lower and upper approximations, we define three binary relations on subsets of U:

$$\begin{aligned} X \approx_* Y &\iff \underline{apr}(X) = \underline{apr}(Y), \\ X \approx^* Y &\iff \overline{apr}(X) = \overline{apr}(Y), \\ X \approx Y &\iff \underline{apr}(X) = \underline{apr}(Y) \text{ and } \overline{apr}(X) = \overline{apr}(Y). \end{aligned} \tag{1.13}$$

Each of them defines an equivalence relation on 2^U, which induces a partition of 2^U. By interpreting an equivalence, say $[X]_{\approx}$ containing X, as a rough set, called a P-rough set, we obtain three algebras of rough sets.

Consider the equivalence relation $\approx$. The set of all P-rough sets is denoted by $R_{\approx}(apr) = 2^U/\approx$. Given two sets $X_1, X_2 \in \mathrm{Com}(apr)$ with $X_1 \subseteq X_2$, if there exists at least a subset $X \subseteq U$ such that $\underline{apr}(X) = X_1$ and $\overline{apr}(X) = X_2$, the following family of subsets of U,

$$\langle X_1, X_2 \rangle = \{X \in 2^U \mid \underline{apr}(X) = X_1, \overline{apr}(X) = X_2\}, \tag{1.14}$$

is called a P-rough set. A set $X \in \langle X_1, X_2 \rangle$ is said to be a member of the P-rough set. Given a member X, a P-rough set can also be more conveniently expression as $[X]_{\approx}$, which is the equivalent class containing X. A member is also referred to as a generator of the P-rough set [3]. Rough set intersection

$\sqcap$, union $\sqcup$, and complement $\neg$ are defined by set operators as follows: for two P-rough sets $\langle X_1, X_2\rangle$ and $\langle Y_1, Y_2\rangle$,

$$\begin{aligned}
&\langle X_1, X_2\rangle \sqcap \langle Y_1, Y_2\rangle \\
&\quad = \{X \in 2^U \mid \underline{apr}(X) = X_1 \cap Y_1, \overline{apr}(X) = X_2 \cap Y_2\} \\
&\quad = \langle X_1 \cap Y_1, X_2 \cap Y_2\rangle, \\
&\langle X_1, X_2\rangle \sqcup \langle Y_1, Y_2\rangle \\
&\quad = \{X \in 2^U \mid \underline{apr}(X) = X_1 \cup Y_1, \overline{apr}(X) = X_2 \cup Y_2\} \\
&\quad = \langle X_1 \cup Y_1, X_2 \cup Y_2\rangle, \\
&\neg\langle X_1, X_2\rangle \\
&\quad = \{X \in 2^U \mid \underline{apr}(X) = \sim X_2, \overline{apr}(X) = \sim X_1\}, \\
&\quad = \langle \sim X_2, \sim X_1\rangle. \qquad (1.15)
\end{aligned}$$

The results are also P-rough sets. The induced system $(R_{\approx}(apr), \sqcap, \sqcup)$ is a complete distributive lattice [1, 30], with zero element $[\emptyset]_{\approx}$ and unit element $[U]_{\approx}$. The corresponding order relation is called P-rough set inclusion given by:

$$\langle X_1, X_2\rangle \sqsubseteq \langle Y_1, Y_2\rangle \iff X_1 \subseteq Y_1 \text{ and } X_2 \subseteq Y_2. \qquad (1.16)$$

The system $(R_{\approx}(apr), \sqcap, \sqcup, \neg, [\emptyset]_{\approx}, [U]_{\approx})$ is called a P-rough set algebra. If equivalence relations $\approx_*$ and $\approx^*$ are used, similar structures can be obtained.

Example 1 This example illustrates the main ideas developed so far. Consider a universe consisting of three elements $U = \{a, b, c\}$ and an equivalence relation $\Re$ on U:

$$a\Re a, \quad b\Re b, \quad b\Re c, \quad c\Re b, \quad c\Re c.$$

The equivalence relation induces two equivalence classes $[a]_{\Re} = \{a\}$ and $[b]_{\Re} = [c]_{\Re} = \{b, c\}$. Table 1 summarizes the lower and upper approximations, the positive, negative and boundary regions, and the accuracy of approximations for all subsets of U. The family of all composed sets is $\mathrm{Com}(apr) = \{\emptyset, \{a\}, \{b, c\}, U\}$. It defines nine I-rough sets. Figure 2 shows the lattice formed by these I-rough sets. Based on the lower and upper approximations, a relation $\approx$ on 2^U is given by:

$$\begin{aligned}
&\emptyset \approx \emptyset, \\
&\{a\} \approx \{a\}, \\
&\{b,c\} \approx \{b,c\}, \\
&\{b\} \approx \{b\}, \quad \{b\} \approx \{c\}, \quad \{c\} \approx \{c\}, \quad \{c\} \approx \{b\}, \\
&\{a,b\} \approx \{a,b\}, \quad \{a,b\} \approx \{a,c\}, \quad \{a,c\} \approx \{a,b\}, \quad \{a,c\} \approx \{a,c\}, \\
&U \approx U.
\end{aligned}$$

X	$\underline{apr}(X)$	$\overline{apr}(X)$	POS(X)	NEG(X)	BND(X)	$\rho(X)$
$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	U	$\emptyset$	1
$\{a\}$	$\{a\}$	$\{a\}$	$\{a\}$	$\{b,c\}$	$\emptyset$	1
$\{b\}$	$\emptyset$	$\{b,c\}$	$\emptyset$	$\{a\}$	$\{b,c\}$	0
$\{c\}$	$\emptyset$	$\{b,c\}$	$\emptyset$	$\{a\}$	$\{b,c\}$	0
$\{a,b\}$	$\{a\}$	U	$\{a\}$	$\emptyset$	$\{b,c\}$	1/3
$\{a,c\}$	$\{a\}$	U	$\{a\}$	$\emptyset$	$\{b,c\}$	1/3
$\{b,c\}$	$\{b,c\}$	$\{b,c\}$	$\{b,c\}$	$\{a\}$	$\emptyset$	1
U	U	U	U	$\emptyset$	$\emptyset$	1

Table 1 Basic notions in Pawlak rough set model

This relation induces the following equivalence classes, i.e., P-rough sets:

$$\begin{aligned}
\langle \emptyset, \emptyset\rangle &= \{\emptyset\}, \\
\langle \{a\}, \{a\}\rangle &= \{\{a\}\}, \\
\langle \emptyset, \{b,c\}, \rangle &= \{\{b\}, \{c\}\}, \\
\langle \{b,c\}, \{b,c\}\rangle &= \{\{b,c\}\}, \\
\langle \{a\}, U\rangle &= \{\{a,b\}, \{a,c\}\}, \\
\langle U, U\rangle &= \{U\}.
\end{aligned}$$

Figure 3 is the lattice formed by these P-rough sets. From this example, one can see that I-rough set algebra is different from the P-rough set algebra. In general, the lattice formed by P-rough sets is isomorphic to a sublattice of the lattice formed by I-rough sets.

2.2 Non-standard rough set models

The Pawlak rough set model may be extended by using an arbitrary binary relation [41, 43]. Given a binary relation $\Re$ and two elements $x, y \in U$, if $x\Re y$, we say that y is $\Re$-related to x. A binary relation may be more conveniently represented by a mapping $r : U \longrightarrow 2^U$:

$$r(x) = \{y \in U \mid x\Re y\}. \qquad (1.17)$$

That is, $r(x)$ consists of all $\Re$-related elements of x. It may be interpreted as a neighborhood of x [12, 14]. If $\Re$ is an equivalence relation, $r(x)$ is the equivalence class containing x. By using the notion of neighborhoods to replace

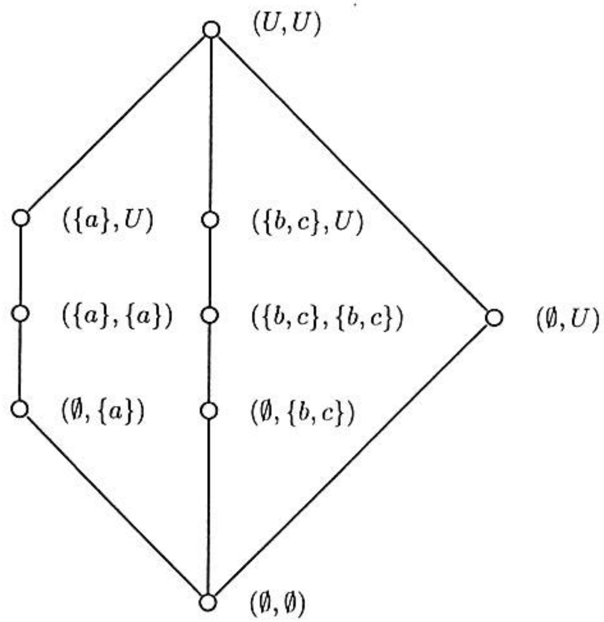

Figure 2 An example of I-rough set algebra

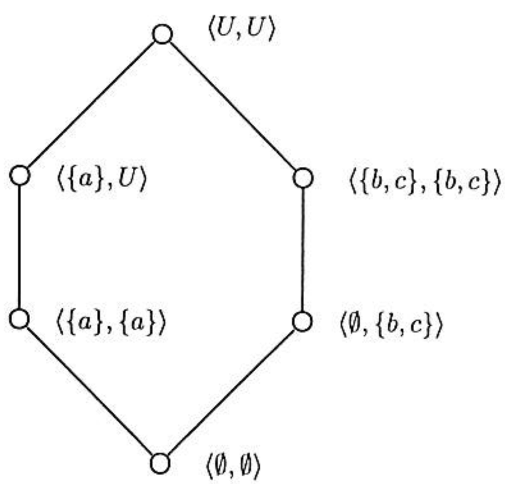

Figure 3 An example of P-rough set algebra

equivalence classes, we can extend equation (1.3) as follows:

$$\begin{aligned} \underline{apr}(X) &= \{x \mid r(x) \subseteq X\}, \\ \overline{apr}(X) &= \{x \mid r(x) \cap X \neq \emptyset\}. \end{aligned} \qquad (1.18)$$

The set $\underline{apr}(X)$ consists of those elements whose $\Re$-related elements are all in X, and $\overline{apr}(X)$ consists of those elements such that at least one of whose $\Re$-related elements is in X. They are referred to as generalized approximations of X.

Generalized approximation operators do not necessarily satisfy all the properties in Pawlak rough set models. Nevertheless, properties (AL1)-(AL5) and (AU1)-(AU5) hold independent of the properties of the binary relation. Properties (AL7)-(AL10) may be used to characterize various rough set models. Such a classification of rough set models is similar to the classification of modal logics. For this purpose, we use the following properties, adopting the same labeling system from Chellas [4]:

(K)	$\underline{apr}(\sim X \cup Y) \subseteq \sim\underline{apr}(X) \cup \underline{apr}(Y)$,
(D)	$\underline{apr}(X) \subseteq \overline{apr}(X)$,
(T)	$\underline{apr}(X) \subseteq X$,
(B)	$X \subseteq \underline{apr}(\overline{apr}(X))$,
(4)	$\underline{apr}(X) \subseteq \underline{apr}(\underline{apr}(X))$,
(5)	$\overline{apr}(X) \subseteq \underline{apr}(\overline{apr}(X))$.

Property (K) does not depend on any particular binary relation. In order to construct a rough set model so that other properties hold, it is necessary to impose certain conditions on the binary relation $\Re$.

Each of the properties (D)-(5) corresponds to a property of the binary relation. Property (D) holds if $\Re$ is a serial relation, i.e., for all $x \in U$, there exists at least an element y such that $x\Re y$. Property (T) holds if $\Re$ is a reflexive relation, i.e., for all $x \in U$, $x\Re x$. Property (B) holds if $\Re$ is a symmetric relation, i.e., for all $x, y \in U$, $x\Re y$ implies $y\Re x$. Property (4) holds if $\Re$ is a transitive relation, i.e., for all $x, y, z \in U$, $x\Re y$ and $y\Re z$ imply $x\Re z$. Property (5) holds if the $\Re$ is an Euclidean relation, i.e., for all $x, y, z \in U$, $x\Re y$ and $x\Re z$ imply $y\Re z$. By combining these properties, one can construct distinct rough set models. Various rough set models are named according to the properties of the binary relation or the properties of the approximation operators. For example, a rough set model constructed from a symmetric relation is referred to as a symmetric rough set model or the KD model. If $\Re$ is reflexive and symmetric, i.e., $\Re$ is

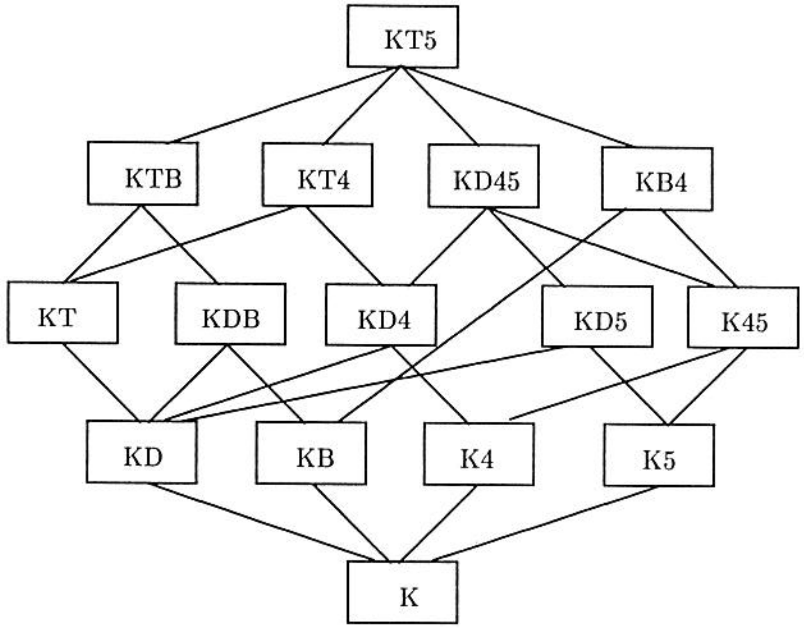

Figure 4 Rough set models

a compatibility relation, properties (K), (D), (T) and (B) hold. This model is labeled by KTB, which is the set of properties satisfied by operators $\underline{apr}$ and $\overline{apr}$. Property (D) does not explicitly appear in the label because it is implied by (T). Similarly, Pawlak rough set model is labled by KT5, which is also commonly known as S5 in modal logic.

By the results from modal logic, it is possible to construct at least fifteen distinct classes of rough set models based on the properties satisfied by the binary relation [4, 17, 43]. Figure 4, adopted from Chellas [4] and Marchal [17], summarizes the relationships between these models. A line connecting two models indicates the model in the upper level is a model in the lower level. These lines that can be derived from the transitivity are not explicitly shown. The model K may be considered as the basic and the weakest model. It does not require any special property on the binary relation. All other models are built on top the model K. The model KT5, i.e., the Pawlak rough set model, is the strongest model.

In the above formulation of rough set model, one considers only two special kinds of relationships between the neighborhood $r(x)$ of an element x and a set X to be approximated. An element belongs to the lower approximation of a subset X if *all* its $\Re$-related elements belong to X, it belongs to the upper approximation if there exists *one* elements belonging to X. The degree of overlap of X and $r(x)$ is not taken into consideration. By employing such information, graded rough set models can be obtained, in the same way graded modal logic is developed [2, 7, 9, 35, 36, 43].

Given the universe U and a binary relation $\Re$ on U, a family of graded approximation operators are defined as:

$$\begin{aligned} \underline{apr}_n(X) &= \{x \mid |r(x)| - |X \cap r(x)| \leq n\}, \\ \overline{apr}_n(X) &= \{x \mid |X \cap r(x)| > n\}. \end{aligned} \tag{1.19}$$

An element of U belongs to $\underline{apr}_n(X)$ if at most n of its $\Re$-related elements are not in X, and belongs to $\overline{apr}_n(X)$ if more than n of its $\Re$-related elements are in X. Based on the properties of binary relation, we can similarly define different classes of graded rough set models.

2.3 Rough sets in information systems

Following Lipski [16], Orlowska [20], Pawlak [21], Vakarelov [34], and Yao and Noroozi [45], we define a set-based information system to be a quadruple,

$$S = (U, At, \{V_a \mid a \in At\}, \{f_a \mid a \in At\}),$$

where

> U is a nonempty set of objects,
> At is a nonempty set of attributes,
> V_a is a nonempty set of values for each attribute $a \in At$,
> $f_a : U \longrightarrow 2^{V_a}$ is an information function for each attribute $a \in At$.

The notion of information systems provides a convenient tool for the representation of objects in terms of their attribute values. If all information functions map an object to only singleton subsets of attribute values, we obtain a degenerate set-based information system commonly used in the Pawlak rough set model. In this case, information functions can be expressed as $f_a : U \longrightarrow V_a$. In the following discussions, we only consider this kind of information systems.

We can describe relationships between objects through their attribute values. With respect to an attribute $a \in At$, a relation $\Re_a$ is given by: for $x, y \in U$,

$$x \Re_a y \iff f_a(x) = f_a(y). \tag{1.20}$$

That is, two objects are considered to be indiscernible, in the view of single attribute a, if and only if they have exactly the *same* value. $\Re_a$ is an equivalence relation. The reflexivity, symmetry and transitivity of $\Re_a$ follow trivially from the properties of the relation $=$ between attribute values. For a subset of attributes $A \subseteq At$, this definition can be extended as follows:

$$x \Re_A y \iff (\forall a \in A) f_a(x) = f_a(y). \tag{1.21}$$

That is, in terms of all attributes in A, x and y are indiscernible, if and only if they have the same value for every attribute in A. The extended relation is still is an equivalence relation [20].

The above discussion provides a convenient and practical method for constructing a binary relation, and in turn a Pawlak rough set model. All other notions can be easily defined. For an element $x \in U$, its equivalence class is given by:

$$r_A(x) = \{y \mid x \Re_A y\}. \tag{1.22}$$

For any subset $X \subseteq U$, the lower and upper approximations can be constructed as:

$$\begin{aligned} \underline{apr}_A(X) &= \{x \mid r_A(x) \subseteq X\}, \\ \overline{apr}_A(X) &= \{x \mid r_A(x) \cap X \neq \emptyset\}. \end{aligned} \tag{1.23}$$

As shown in the following example, different subsets of attributes may induce distinct approximation space, and hence different approximations of the same set.

Example 2 Consider the information system given in Table 2, taken from Quinlan [31]. Each object is described by three attributes. If the attribute $A = \{\text{Hair}\}$ is chosen, we can partition the universe into equivalence classes $\{o_1, o_2, o_6, o_8\}$, $\{o_3\}$, and $\{o_4, o_5, o_7\}$, reflecting the colour of Hair being blond, red and dark, respectively. With respect to the class $+ = \{o1, o_3, o_6\}$, the following approximations are obtained:

$$\begin{aligned} \underline{apr}_A(+) &= \{o_3\}, \\ \overline{apr}_A(+) &= \{o_1, o_2, o_3, o_6, o_8\}. \end{aligned}$$

Object	Height	Hair	Eyes	Classification
o_1	short	blond	blue	+
o_2	short	blond	brown	-
o_3	tall	red	blue	+
o_4	tall	dark	blue	-
o_5	tall	dark	blue	-
o_6	tall	blond	blue	+
o_7	tall	dark	brown	-
o_8	short	blond	brown	-

Table 2 An information system

Hence,

$$\begin{aligned} \text{POS}_A(+) &= \underline{apr}_A(+) = \{o_3\}, \\ \text{BND}_A(+) &= \overline{apr}_A(+) - \underline{apr}_A(+) = \{o_1, o_2, o_6, o_8\}, \\ \text{NEG}_A(+) &= U - \overline{apr}_A(+) = \{o_4, o_5, o_7\}. \end{aligned}$$

If a set of two attributes $A' = \{\text{Hair}, \text{Eyes}\}$ is used, we have equivalence classes $\{o_1, o_6\}$, $\{o_2, o_8\}$, $\{o_3\}$, $\{o_4, o_5\}$ and $\{o_7\}$. The lower and upper approximation of $+$ are:

$$\begin{aligned} \underline{apr}_{A'}(+) &= \{o_1, o_3, o_6\}, \\ \overline{apr}_{A'}(+) &= \{o_1, o_3, o_6\}. \end{aligned}$$

Three regions are:

$$\begin{aligned} \text{POS}_{A'}(+) &= \underline{apr}_{A'}(+) = \{o_1, o_3, o_6\}, \\ \text{BND}_{A'}(+) &= \overline{apr}_{A'}(+) - \underline{apr}_{A'}(+) = \emptyset, \\ \text{NEG}_{A'}(+) &= U - \overline{apr}_{A'}(+) = \{o_2, o_4, o_5, o_7, o_8\}. \end{aligned}$$

From this example, it is clear that some approximation spaces are better than others.

The Pawlak rough set model can be easily generalized in information system by considering any binary relations on attribute values, instead of the trivial equality relation $=$. Suppose R_a is a binary relation on the values of an attribute $a \in At$. By extending equation (1.20), for $a \in At$ we define a binary relation on U:

$$x \Re_a y \iff f_a(x) R_a f_a(y). \tag{1.24}$$

Similarly, by extending equation (1.21), for $A \subseteq At$ we define a relation on U:

$$\begin{aligned} x\Re_A y &\iff (\forall a \in A) f_a(x) R_a f_a(y) \\ &\iff (\forall a \in A) x \Re_a y. \end{aligned} \tag{1.25}$$

An object x is related to another object y, based on an attribute a, if their values on a are related. With respect to a subset A of attributes, x is related to y if their values are related for every attribute in A. When all relations R_a are chosen to be $=$, the proposed definition reduced to the definition in the Pawlak rough set model.

The empty set $\emptyset$ produces the coarsest relation, i.e., $\Re_\emptyset = U \times U$, where $\times$ denotes the Cartesian product of sets. If the entire attribute set is used, one obtains the finest relation $\Re_{At}$. Moreover, if each object is described by an unique description, $\Re_{At}$ becomes the identity relation. The algebra $(\{\Re_A\}_{A \subseteq At}, \cap)$ is a lower semilattice with the zero element $\Re_{At}$ [20].

The relation $\Re_a$ preserves properties of R_a. Suppose R_a is a binary relation on V_a, and $\Re_a$ a binary relation on U defined by equation (1.24). Then,

a). R_a is serial $\Longrightarrow \Re_a$ is serial;
b). R_a is reflexive $\Longrightarrow \Re_a$ is reflexive;
c). R_a is symmetric $\Longrightarrow \Re_a$ is symmetric;
d). R_a is transitive $\Longrightarrow \Re_a$ is transitive;
e). R_a is Euclidean $\Longrightarrow \Re_a$ is Euclidean.

The set of $\Re_A$-related objects, $r_A(x) = \{y \mid x\Re_A y\}$, can be regarded as a neighborhood of x. Likewise, the set of R_a-related values, $r_a(v) = \{v' \mid v R_a v'\}$, can be viewed as a neighborhood of v [12]. By definition, a neighborhood of objects is defined according to neighborhoods of its attribute values:

$$\begin{aligned} r_A(x) &= \{y \mid x\Re_A y\} \\ &= \bigcap_{a \in A} \{y \mid x \Re_a y\} \\ &= \bigcap_{a \in A} \{y \mid f_a(x) R_a f_a(y)\} \\ &= \bigcap_{a \in A} \{y \mid f_a(y) \in r_a(f_a(x))\}. \end{aligned} \tag{1.26}$$

This suggests that the notion of generalized rough sets is useful for approximate retrieval in information systems.

2.4 Rough set model over two universes

Recently, Wong, Wang and Yao generalized the rough set model using two distinct but related universes [38, 39, 47]. Let U and V represent two finite universes of interest. Suppose the relationships between elements of the two universes are described by a compatibility relation [32]. The formulation and interpretation of U and V and the compatibility relation between the two universes depend very much on the available knowledge and the domain of applications. For example, in a medical diagnosis system, U can be a set of symptoms and V a set of diseases. A symptom $u \in U$ is said to be compatible with a disease $v \in V$ if any patient with symptom u may have contracted the disease v. An element $u \in U$ is compatible with an element $v \in V$, written $u \, \mathcal{C} \, v$, if the u is related to v. Without loss of generality, we may assume that for any $u \in U$ there exists a $v \in V$ with $u \, \mathcal{C} \, v$, and vice versa.

A compatibility relation $\mathcal{C}$ between U and V can be equivalently defined by a multi-valued mapping, $\gamma : U \longrightarrow 2^V$, as [5, 32]:

$$\gamma(u) = \{v \in V \mid u \, \mathcal{C} \, v\}. \tag{1.27}$$

That is, $\gamma(u)$ is a subset of V consisting of all elements compatible with u. Based on this multi-valued mapping, a subset $X \subseteq V$ may be represented in terms of these elements of U compatible with the elements in X. For example a particular group of diseases may be described by the symptoms compatible with them. Since the induced multi-valued mapping is not necessarily an one-to-one mapping, one may not be able to derive an exact representation for any subset $X \subseteq V$. By extending notion of approximation operators in rough set model, we define a pair of lower and upper approximations:

$$\begin{aligned} \underline{apr}(X) &= \{u \in U \mid \gamma(u) \subseteq X\}, \\ \overline{apr}(X) &= \{u \in U \mid \gamma(u) \cap X \neq \emptyset\}. \end{aligned} \tag{1.28}$$

The set $\underline{apr}(X)$ consists of the elements in U compatible with only those elements in X, while the set $\overline{apr}(X)$ consists of the elements in U compatible with at least one element in X. Therefore, the lower approximation $\underline{apr}(X)$ can be interpreted as the pessimistic description and the upper approximation $\overline{apr}(X)$ as the optimistic description of X. These approximation operators satisfy properties similar to (AL1)-(AL6) and (AU1)-(AU6). Since two universes are involved, there do not exist properties similar to (AL7)-(AL10) and (AU7)-(AU10).

2.5 Rough set model over Boolean algebras

Recall that the set of composed sets Com(apr) forms a sub-algebra of the Boolean algebra of the power set. One can easily formulate Pawlak rough set model in a wider context of Boolean algebra. Suppose $(\mathcal{A}, \wedge, \vee, \neg, 0, 1)$ is a Boolean algebra and $(\mathcal{B}, \wedge, \vee, \neg, 0, 1)$ is a sub-algebra. In terms of elements of $\mathcal{B}$, one may approximate any element of $\mathcal{A}$ using a pair of lower and upper approximations: for $a \in \mathcal{A}$,

$$\begin{aligned} \underline{apr}(a) &= \bigvee\{b \mid b \in \mathcal{B}, b \preceq a\}, \\ \overline{apr}(a) &= \bigwedge\{b \mid b \in \mathcal{B}, a \preceq b\}. \end{aligned} \tag{1.29}$$

Clearly, this definition reduce to Pawlak's original proposal if $\mathcal{A}$ is chosen to be 2^U and $\mathcal{B}$ is chosen to be Com(apr).

Wong, Wang and Yao [39] extended the above formulation further by considering two arbitrary Boolean algebras. Suppose $\underline{f}\colon \mathcal{A} \longrightarrow \mathcal{B}$ and $\overline{f}\colon \mathcal{A} \longrightarrow \mathcal{B}$ are two mappings from a Boolean algebra $(\mathcal{A}, \vee, \wedge, \neg, 0, 1)$ to another Boolean algebra $(\mathcal{B}, \vee, \wedge, \neg, 0, 1)$. We say that $\underline{f}$ and $\overline{f}$ are dual mappings if $\overline{f}(a) = \neg\underline{f}(\neg a)$ for every $a \in \mathcal{A}$. The pair of dual mappings form an interval structure if they satisfy the following axioms:

$$\begin{array}{ll} \text{(IL1)} & \underline{f}(a) \wedge \underline{f}(b) = \underline{f}(a \wedge b), \\ \text{(IL2)} & \underline{f}(0) = 0, \\ \text{(IL3)} & \underline{f}(1) = 1, \\ \text{(IU1)} & \overline{f}(a \vee b) = \overline{f}(a) \vee \overline{f}(b), \\ \text{(IU2)} & \overline{f}(0) = 0, \\ \text{(IU3)} & \overline{f}(1) = 1. \end{array}$$

These properties indeed correspond to properties (AL3), (AL2), (AL6), (AU3), (AU2) and (AU6).

An alternate way of defining an interval structure is through another mapping $j\colon \mathcal{A} \longrightarrow \mathcal{B}$ satisfying the axioms:

$$\begin{array}{ll} \text{(A1)} & j(0) \;=\; 0, \\ \text{(A2)} & \bigvee_{a \in \mathcal{A}} j(a) = 1, \\ \text{(A3)} & a \neq b \Longrightarrow j(a) \wedge j(b) = 0. \end{array}$$

This mapping is called a *basic assignment*, and an element $a \in \mathcal{A}$ with $j(a) \neq 0$ is called a *focal element*. From a given j, one can define a mapping $\underline{f}$: for all $a \in \mathcal{A}$,

$$\underline{f}(a) \;=\; \bigvee_{b \preceq a} j(b), \tag{1.30}$$

and another mapping $\overline{f}$ by the relationship $\overline{f}(a) = \neg\underline{f}(\neg a)$. The mapping $\overline{f}$ can be equivalently defined by:

$$\overline{f}(a) \;=\; \bigvee_{a \wedge b \neq 0} j(b). \tag{1.31}$$

Conversely, given an interval structure $(\underline{f}, \overline{f})$, we can construct the basic assignment j by the formula: for all $a \in \mathcal{A}$,

$$j(a) = \underline{f}(a) \wedge \neg(\bigvee_{b \prec a} \underline{f}(b)). \tag{1.32}$$

Rough set models on the same universe and on two universes are only special cases of this general framework. Based on the axioms of an interval structure, the above developed relationships hold in any rough set model that is stronger than the KD model. More specifically, we have the following connections:

$$\begin{aligned} j(X) &= \{x \mid r(x) = X\}, \\ j(X) &= \underline{apr}(X) - \bigcap_{Y \subset X} \underline{apr}(Y), \\ \underline{apr}(X) &= \bigcup_{Y \subseteq X} j(X), \\ \overline{apr}(X) &= \bigcup_{Y \cap X \neq \emptyset} j(X). \end{aligned} \tag{1.33}$$

Therefore, the basic assignment provides another representation of approximation operators.

3 PROBABILISTIC ROUGH SET MODELS

Based on the notion of rough membership functions, we review two different approaches for the construction of probabilistic rough set model. One is related to probabilistic modal logic and the other is based on decision theory.

3.1 Rough membership functions

Pawlak and Skowron [28], Pawlak *et al.* [29] and Wong and Ziarko [40] proposed another way to characterize a rough set by a single membership function. For any $X \subseteq U$, a rough membership function is defined by:

$$\mu_X(x) = \frac{|X \cap [x]_{\Re}|}{|[x]_{\Re}|}. \tag{1.34}$$

By definition, elements in the same equivalent class have the same degree of membership. The rough membership $\mu_X(x)$ may be interpreted as the probability of x belonging to X given that x belongs to an equivalence class. This interpretation leads to probabilistic rough sets [29, 40]. Like the algebraic rough set model, the intersection and union of probabilistic rough sets are not truth-functional. Nevertheless, we have:

(m1) $\mu_X(x) = 1 \Longleftrightarrow x \in \mathrm{POS}(X)$,
(m2) $\mu_X(x) = 0 \Longleftrightarrow x \in \mathrm{NEG}(X)$,
(m3) $0 < \mu_X(x) < 0 \Longleftrightarrow x \in \mathrm{BND}(X)$,
(m4) $\mu_{\sim X}(x) = 1 - \mu_X(x)$,
(m5) $\mu_{X\cup Y}(x) = \mu_X(x) + \mu_Y(x) - \mu_{X\cap Y}(x)$,
(m6) $\max(0, \mu_X(x) + \mu_Y(x) - 1) \leq \mu_{X\cap Y}(x) \leq \min(\mu_X(x), \mu_Y(x))$,
(m7) $\max(\mu_X(x), \mu_Y(x)) \leq \mu_{X\cup Y}(x) \leq \min(1, \mu_X(x) + \mu_Y(x))$.

They follow from the property of probability. The definition in equation (1.34) can be easily extended by using an arbitrary binary relation.

3.2 Variable precision rough set model

In the definition of graded rough set models, the size of $r(x)$ is not taken into consideration. By using such information, we can define variable precision, or probabilistic, rough set model [40, 49], in parallel to probabilistic modal logic [8, 10, 19, 43].

With respect to the universe U and a binary relation $\Re$ on U, we define a family of probabilistic rough set operators:

$$\begin{aligned} \underline{apr}_{\alpha}(X) &= \{x \mid \frac{|X \cap r(x)|}{|r(x)|} \geq 1 - \alpha\}, \\ \overline{apr}_{\alpha}(X) &= \{x \mid \frac{|X \cap r(x)|}{|r(x)|} > \alpha\}. \end{aligned} \tag{1.35}$$

By definition, for a serial binary relation and $\alpha \in [0,1]$, probabilistic rough set operators satisfy the following properties:

(PL0) $\underline{apr}(X) = \underline{apr}_0(X)$,
(PL1) $\underline{apr}_{\alpha}(X) = \sim\overline{apr}_{\alpha}(\sim X)$,
(PL2) $\underline{apr}_{\alpha}(U) = U$,
(PL3) $\underline{apr}_{\alpha}(X \cap Y) \subseteq \underline{apr}_{\alpha}(X) \cap \underline{apr}_{\alpha}(Y)$,
(PL4) $\underline{apr}_{\alpha}(X \cup Y) \supseteq \underline{apr}_{\alpha}(X) \cup \underline{apr}_{\alpha}(Y)$,
(PL5) $X \subseteq Y \Longrightarrow \underline{apr}_{\alpha}(X) \subseteq \underline{apr}_{\alpha}(Y)$,
(PL6) $\alpha \geq \beta \Longrightarrow \underline{apr}_{\alpha}(X) \supseteq \underline{apr}_{\beta}(X)$,
(PU0) $\overline{apr}(X) = \overline{apr}_0(X)$,
(PU1) $\overline{apr}_{\alpha}(X) = \sim\underline{apr}_{\alpha}(\sim X)$,
(PU2) $\overline{apr}_{\alpha}(\emptyset) = \emptyset$,
(PU3) $\overline{apr}_{\alpha}(X \cup Y) \supseteq \overline{apr}_{\alpha}(X) \cup \overline{apr}_{\alpha}(Y)$,
(PU4) $\overline{apr}_{\alpha}(X \cap Y) \subseteq \overline{apr}_{\alpha}(X) \cap \overline{apr}_{\alpha}(Y)$,
(PU5) $X \subseteq Y \Longrightarrow \overline{apr}_{\alpha}(X) \subseteq \overline{apr}_{\alpha}(Y)$,
(PU6) $\alpha \geq \beta \Longrightarrow \overline{apr}_{\alpha}(X) \subseteq \overline{apr}_{\beta}(X)$.

Moreover, for $0 \leq \alpha < 0.5$,

(PD) $\underline{apr}_{\alpha}(X) \subseteq \overline{apr}_{\alpha}(X)$,

which may be interpreted as a probabilistic version of axiom (D). In this case, one can also partition the into three regions based on the value of α:

$$\begin{aligned} \mathrm{POS}_{\alpha}(X) &= \underline{apr}_{\alpha}(X), \\ \mathrm{NEG}_{\alpha}(X) &= U - \overline{apr}_{\alpha}(X), \\ \mathrm{BND}_{\alpha}(X) &= \overline{apr}_{\alpha}(X) - \underline{apr}_{\alpha}(X). \end{aligned} \tag{1.36}$$

They may be referred to as the probabilistic positive, negative and boundary regions. In the following subsection, we will show that the value of α can be determined within the framework of decision theory.

3.3 Rough set model based on decision theory

In the variable precision rough set model, the universe is partitioned into three regions. The same goal can be achieved by using rough membership functions in the framework of decision theory [46]. In terms of decision-theoretic language we have a set of states $\Omega = \{X, \neg X\}$, indicating that an element belongs to and does not belong to X, and the set of actions $A = \{a_1, a_2, a_3\}$, representing the three actions, deciding $\text{POS}(X)$, deciding $\text{NEG}(X)$, and deciding $\text{BND}(X)$, respectively.

Let $\lambda(a_i|X)$ denote the loss incurred for taking action a_i when an object in fact belongs to X, and let $\lambda(a_i|\neg X)$ denote the loss incurred when the object actually belongs to $\neg X$. $P(X \mid r(x))$ and $P(\neg X \mid r(x))$ are the probabilities that an object with neighborhood $r(x)$ belongs to X and $\neg X$, respectively. They are in fact the rough membership functions with respect to X and $\neg X$. Thus, the expected loss $R(a_i|r(x))$ associated with taking the individual actions can be expressed as:

$$\begin{aligned} R(a_1|r(x)) &= \lambda_{11}P(X \mid r(x)) + \lambda_{12}P(\neg X \mid r(x)), \\ R(a_2|r(x)) &= \lambda_{21}P(X \mid r(x)) + \lambda_{22}P(\neg X \mid r(x)), \\ R(a_3|r(x)) &= \lambda_{31}P(X \mid r(x)) + \lambda_{32}P(\neg X \mid r(x)), \end{aligned} \tag{1.37}$$

where $\lambda_{i1} = \lambda(a_i|X)$, $\lambda_{i2} = \lambda(a_i|\neg X)$, and $i = 1, 2, 3$. The Bayesian decision procedure leads to the following minimum-risk decision rules:

(P) Decide $\text{POS}(X)$ if
$$R(a_1|r(x)) \leq R(a_2|r(x)) \text{ and } R(a_1|r(x)) \leq R(a_3|r(x));$$
(N) Decide $\text{NEG}(X)$ if
$$R(a_2|r(x)) \leq R(a_1|r(x)) \text{ and } R(a_2|r(x)) \leq R(a_3|r(x));$$
(B) Decide $\text{BND}(X)$ if
$$R(a_3|r(x)) \leq R(a_1|r(x)) \text{ and } R(a_3|r(x)) \leq R(a_2|r(x)).$$

Since $P(X \mid r(x)) + P(\neg X \mid r(x)) = 1$, the above decision rules can be simplified so that only the probabilities $P(X \mid r(x))$ are involved. Thus, we can classify any object with neighborhood $r(x)$ based only on the probabilities $P(X \mid r(x))$, i.e., the rough membership function, and the given loss function λ_{ij} ($i = 1, 2, 3$; $j = 1, 2$).

Consider a special kind of loss functions with $\lambda_{11} \leq \lambda_{31} < \lambda_{21}$ and $\lambda_{22} \leq \lambda_{32} < \lambda_{12}$. The loss of classifying an object x belonging to X into the positive region $\text{POS}(X)$ is less than or equal to the loss of classifying x into the boundary region $\text{BND}(X)$, and both of these losses are strictly less than the loss of classifying x into the negative region $\text{NEG}(X)$. We obtain the reverse order of losses by classifying an object that does not belong to X. For this type of loss functions, the minimum-risk decision rules (P)-(B) can be written as:

(P) Decide $\text{POS}(X)$ if $P(X \mid r(x)) \geq \beta$ and $P(X \mid r(x)) \geq \gamma$;

(N) Decide $\text{NEG}(X)$ if $P(X \mid r(x)) \leq \gamma$ and $P(X \mid r(x)) \leq \delta$;

(B) Decide $\text{BND}(X)$ if $\delta \leq P(X \mid r(x))$ and $P(X \mid r(x)) \leq \beta$;

where

$$\begin{aligned} \beta &= \frac{\lambda_{12} - \lambda_{32}}{(\lambda_{31} - \lambda_{11}) + (\lambda_{12} - \lambda_{32})}, \\ \gamma &= \frac{\lambda_{12} - \lambda_{22}}{(\lambda_{21} - \lambda_{11}) + (\lambda_{12} - \lambda_{22})}, \\ \delta &= \frac{\lambda_{32} - \lambda_{22}}{(\lambda_{21} - \lambda_{31}) + (\lambda_{32} - \lambda_{22})} \end{aligned} \tag{1.38}$$

From the assumptions, $\lambda_{11} \leq \lambda_{31} < \lambda_{21}$ and $\lambda_{22} \leq \lambda_{32} < \lambda_{12}$, it follows that $\beta \in (0, 1]$, $\gamma \in (0, 1)$, and $\delta \in [0, 1)$. Decision rules (P)-(B) depend only on the parameters β, γ, and δ computable from the λ_{ij}'s directly supplied by the user.

If $\delta \leq \beta$, $\delta \leq \gamma \leq \beta$. By decision rules (P)-(B), three regions can be determined by δ and β. If $\beta < \delta$, we have $\beta < \gamma < \delta$. According to (P)-(B), the boundary region is empty, and both positive and negative region can be determined by γ. To be consistent with the variable precision rough set model, we assume $\delta < \beta$, which implies $\delta < \gamma < \beta$. Furthermore, we choose a tie-breaking rule to differentiate actions producing the same risk. If the risk of deciding $\text{POS}(X)$ or $\text{BND}(X)$ is the same, we decide $\text{POS}(X)$; if the risk of deciding $\text{NEG}(X)$ or $\text{BND}(X)$ is the same, we decide $\text{NEG}(X)$. Under these assumptions, (P)-(B) can be simplified into:

(P) Decide $\text{POS}(X)$ if $P(X \mid r(x)) \geq \beta$;

(N) Decide $\text{NEG}(X)$ if $P(X \mid r(x)) \leq \delta$;

(B) Decide $\text{BND}(X)$ if $\delta < P(X \mid r(x)) < \beta$.

The positive, negative, and boundary regions can be explicitly expressed in terms of the pair of parameters δ and β, namely:

$$\begin{aligned} \text{POS}_{\beta,\delta}(X) &= \{x \mid P(X \mid r(x)) \geq \beta\}, \\ \text{NEG}_{\beta,\delta}(X) &= \{x \mid P(X \mid r(x)) \leq \delta\}, \\ \text{BND}_{\beta,\delta}(X) &= \{x \mid \delta < P(X \mid r(x)) < \beta\}. \end{aligned} \tag{1.39}$$

The lower and upper approximations $\underline{apr}_{\beta,\delta}(X)$ and $\overline{apr}_{\beta,\delta}(X)$ of X can be defined as:

$$\begin{aligned} \underline{apr}_{\beta,\delta}(X) &= \text{POS}_{\beta,\delta}(X) \\ &= \{x \mid P(X \mid r(x)) \geq \beta\}, \\ \overline{apr}_{\beta,\delta}(X) &= \text{POS}_{\beta,\delta}(X) \cup \text{BND}_{\beta,\delta}(X) \\ &= \{x \mid P(X \mid r(x)) > \delta\}. \end{aligned} \tag{1.40}$$

Now assume the following condition:

$$\frac{\lambda_{12} - \lambda_{32}}{\lambda_{31} - \lambda_{11}} = \frac{\lambda_{21} - \lambda_{31}}{\lambda_{32} - \lambda_{22}}. \quad (1.41)$$

We have $\beta = 1 - \delta$. Let $\alpha = \delta$. The lower and upper approximations can be expressed by:

$$\begin{aligned} \underline{apr}_{1-\alpha,\alpha}(X) &= \{x \mid P(X \mid r(x)) \geq 1 - \alpha\}, \\ \overline{apr}_{1-\alpha,\alpha}(X) &= \{x \mid P(X \mid r(x)) > \alpha\}. \end{aligned} \quad (1.42)$$

They are exactly the probabilistic approximations given in equation (1.35) if the required probabilities are estimated from the cardinalities of $X \cap r(x)$ and $r(x)$, namely, $P(X \mid r(x)) = |X \cap r(x)|/|r(x)|$. The approximations of in an algebraic rough set model can be easily derived. Consider the following loss function:

$$\lambda_{12} = \lambda_{21} = 1, \quad \lambda_{11} = \lambda_{22} = \lambda_{31} = \lambda_{32} = 0. \quad (1.43)$$

This means that there is a unit cost if an object belonging to X is classified into the negative region or if an object not belonging to X is classified into the positive region; otherwise there is no cost. For such a loss function, we obtain from equation (1.38) that $\beta = 1$ and $\delta = 0$. Hence, according to equation (1.40), we have:

$$\begin{aligned} \underline{apr}_{1,0}(X) &= \{x \mid P(X \mid r(x)) = 1\}, \\ \overline{apr}_{1,0}(X) &= \{x \mid P(X \mid r(x)) > 0\}. \end{aligned} \quad (1.44)$$

With the probabilities estimated by

$$P(X \mid r(x)) = \frac{|X \cap r(x)|}{|r(x)|}, \quad (1.45)$$

$\underline{apr}_{1,0}(X)$ and $\overline{apr}_{1,0}(X)$ can be expressed as:

$$\begin{aligned} \underline{apr}_{1,0}(X) &= \{x \mid r(x) \subseteq X\}, \\ \overline{apr}_{1,0}(X) &= \{x \mid r(x) \cap X \neq \emptyset\}. \end{aligned} \quad (1.46)$$

The results given here suggest that both algebraic rough set and probabilistic rough set models can be viewed as a special case of the decision theoretic framework.

4 CONCLUSION

In the Pawlak rough set model, an equivalent relation is used to define an approximation space. Following the argument of Pawlak and using an arbitrary binary relation, one can derive various type of generalized rough set models. Alternatively, one may also generalize Pawlak rough set model by using statistical information. Based on the properties of binary relation, one can identify the properties of lower and upper approximations. Generalized rough set models may be grouped into two classes, the algebraic and probabilistic rough set models, depending on whether statistical information is used. The algebraic class includes normal rough set models, graded rough set models, rough set models over two universes, and rough set models over Boolean algebras. The probabilistic rough set models may be interpreted based on rough membership functions.

The successful applications of the theory of rough sets depends to a large extent on the formulation, characterization, and interpretation of the theory. In this paper, existing works are reviewed using a very simple, and unified, view. That is, rough set models are constructed, classified, and interpreted based on the notion of binary relations. This view in may be useful in the applications of the theory of rough sets.

REFERENCES

[1] Bonikowski, Z., "Algebraic structures of rough sets," in: Rough Sets, Fuzzy Sets and Knowledge Discovery, edited by W.P. Ziarko, Springer-Verlag, London, pp. 242-247, 1994.

[2] de Caro, F., "Graded modalities II," Studia Logica, **XLVII**, pp. 1-10, 1988.

[3] Chanas, S. and Kuchta, D., "Further remarks on the relation between rough and fuzzy sets," Fuzzy Sets and Systems, **47**, pp. 391-394, 1992.

[4] Chellas, B.F., Modal Logic: An Introduction, Cambridge University Press, Cambridge, 1980.

[5] Dempster, A.P., "Upper and lower probabilities induced by a multivalued mapping," Annals of Mathematical Statistics, **38**, pp. 325-339, 1967.

[6] Dubois, D. and Prade, H. "Rough fuzzy sets and fuzzy rough sets," it International Journal of General Systems, **17**, pp. 191-209, 1990.

[7] Fattorosi-Barnaba, M. and de Caro, F., "Graded modalities I," Studia Logica, **XLIV**, pp. 197-221, 1985.

[8] Fattorosi-Barnaba, M. and Amati, G., "Modal operators with probabilistic interpretations I," Studia Logica, **XLVI**, pp. 383-393, 1987.

[9] Fattorosi-Barnaba, M. and de Caro, F., "Graded modalities III," Studia Logica, **XLVII**, pp. 99-110, 1988.

[10] Hart, W.D., "Probability as degree of possibility," Notre-Dame Journal of Formal Logic, **13**, pp. 286-288, 1972.

[11] Iwinski, T.B., "Algebraic approach to rough sets," Bulletin of the Polish Academy of Sciences, Mathematics, **35**, pp. 673-683, 1987.

[12] Lin, T.Y., "Neighborhood systems and approximation in database and knowledge base systems," Proceedings of the Fourth International Symposium on Methodologies of Intelligent Systems, 1989.

[13] Lin, T.Y. (Ed.) Proceedings of the CSC'95 Workshop on Rough Sets and Database Mining, San Jose State University, 1995.

[14] Lin, T.Y. and Liu, Q., "Rough approximate operators: axiomatic rough set theory," in: Rough Sets, Fuzzy Sets and Knowledge Discovery, edited by W.P. Ziarko, Springer-Verlag, London, pp. 256-260, 1994.

[15] Lin, T.Y. and Wildberger, A.M. (Ed.) Soft Computing, The Society for Computer Simulation, San Diego, 1995.

[16] Lipski, W. Jr., "On databases with incomplete information," Journal of the ACM, **28**, 41-70, 1981.

[17] Marchal, B. "Modal logic – a brief tutorial," in: Non-Standard Logics for Automated Reasoning, edited by Smets, P., Mamdani, A., Dubois, D. and Prade, H., New York, Academic Press, pp. 15-23, 1988.

[18] Marczewski, E. and Steinhaus, H., "On a certain distance of sets and the corresponding distance of functions," Colloquium Mathemmaticum, **6**, pp. 319-327, 1958.

[19] Murai, T., Miyakoshi, M. and Shimbo, M., "Fuzzification of modal operators from the standpoint of fuzzy semantics," Proceedings of the Second International Fuzzy Systems Association Congress, pp. 430-433, 1987.

[20] Orlowska, E., "Logic of indiscernibility relations," Bulletin of the Polish Academy of Sciences, Mathematics, **33**, pp. 475-485, 1985.

[21] Pawlak, Z., "Information systems – theoretical foundations," Information Systems, **6**, pp. 205-218, 1981.

[22] Pawlak, Z., "Rough sets," International Journal of Computer and Information Sciences, **11**, pp. 341-356, 1982.

[23] Pawlak, Z. "Rough classification," International Journal of Man-Machine Studies, **20**, pp. 469-483, 1984.

[24] Pawlak, Z. "Rough sets and fuzzy sets," Fuzzy Sets and Systems, **17**, pp. 99-102, 1985.

[25] Pawlak, Z. Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Boston, 1991.

[26] Pawlak, Z. "Rough sets: a new approach to vagueness," in: Fuzzy Logic for the Management of Uncertainty, edited by L.A. Zadeh and J. Kacprzyk, Eds., John Wiley & Sons, New York, pp. 105-118, 1992.

[27] Pawlak, Z. "Hard and soft sets," in: Rough Sets, Fuzzy Sets and Knowledge Discovery, edited by W.P. Ziarko, Springer-Verlag, London, pp. 130-135, 1994.

[28] Pawlak, Z. and Skowron, A., "Rough membership functions," in: Fuzzy Logic for the Management of Uncertainty, edited by L.A. Zadeh and J. Kacprzyk, John Wiley & Sons, New York, pp. 251-271, 1994.

[29] Pawlak, Z., Wong, S.K.M. and Ziarko, W., "Rough sets: probabilistic versus deterministic approach," International Journal of Man-machine Studies, **29**, pp. 81-95, 1988.

[30] Pomykala, J. and Pomykala, J.A., "The Stone algebra of rough sets," Bulletin of the Polish Academy of Sciences, Mathematics, **36**, pp. 495-508, 1988.

[31] Quinlan, J.R., "Learning efficient classification procedures and their application to chess endgames," in: Machine Learning: An Artificial Intelligence Approach, vol. 1, edited by Michalski, J.S., Carbonell, J.G., and Mirchell, T.M., Morgan Kaufmann, Palo Alto, CA, pp. 463-482, 1983.

[32] Shafer, G., "Belief functions and possibility measures," in: Analysis of Fuzzy Information, vol. 1: Mathematics and Logic, edited by Bezdek, J.C., CRC Press, Boca Raton, pp. 51-84, 1987.

[33] Slowinski, R. (Ed.) Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, Boston, 1992.

[34] Vakarelov, D., "A modal logic for similarity relations in Pawlak knowledge representation systems," Fundamenta Informaticae, **XV**, pp. 61-79, 1991.

[35] van der Hoek, W., "On the semantics of graded modalities," Journal of Applied Non-classical Logic, **2**, pp. 81-123, 1992.

[36] van der Hoek, W. and Meyer, J.-J.Ch., "Graded modalities in epistemic logic," Logique et Analyse, **133-134**, pp. 251-270, 1991.

[37] Wasilewska, A. Topological rough algebras. Manuscript, 1995.

[38] Wong, S.K.M., Wang, L.S. and Yao, Y.Y., "Interval structure: a framework for representing uncertain information," Uncertainty in Artificial Intelligence: Proceedings of the 8th Conference, pp. 336-343, 1993.

[39] Wong, S.K.M. and Wang, L.S. and Yao, Y.Y., "On modeling uncertainty with interval structures," Computational Intelligence, **11**, pp. 406-426, 1995.

[40] Wong, S.K.M. and Ziarko, W., "Comparison of the probabilistic approximate classification and the fuzzy set model," Fuzzy Sets and Systems, **21**, pp. 357-362, 1987.

[41] Wybraniec-Skardowska, U., "On a generalization of approximation space," Bulletin of the Polish Academy of Sciences, Mathematics, **37**, pp. 51-61, 1989.

[42] Yao, Y.Y. "On combining rough and fuzzy sets," Proceedings of the CSC'95 Workshop on Rough Sets and Database Mining, edited by Lin, T.Y., San Jose State University, 1995.

[43] Yao, Y.Y. and Lin, T.Y., "Generalization of rough sets using modal logic," Intelligent Automation and Soft Computing, An International Journal, to appear.

[44] Yao, Y.Y., and Lin, T.Y., "Approximate reasoning using rough-set theory," Proceedings of International Symposia on Soft Computing and Intelligent Industrial Automation, pp., 1996.

[45] Yao, Y.Y. and Noroozi, N., "A unified model for set-based computations," in: Soft Computing, edited by Lin, T.Y. and Wildberger, A.M., The Society for Computer Simulation, San Diego, pp. 252-255, 1995.

[46] Yao, Y.Y. and Wong, S.K.M., "A decision theoretic framework for approximating concepts," International Journal of Man-machine Studies, **37**, pp. 793-809, 1992.

[47] Yao, Y.Y., Wong, S.K.M. and Wang, L.S., "A non-numeric approach to uncertain reasoning," International Journal of General Systems, **23**, pp. 343-359, 1995.

[48] Zakowski, W. "Approximations in the space (U, Π)," Demonstratio Mathematica, **XVI**, pp. 761-769, 1983.

[49] Ziarko, W., "Variable precision rough set model," Journal of Computer and System Sciences, **46**, pp. 39-59, 1993.

[50] Ziarko, W.P. (Ed.) Rough Sets, Fuzzy Sets and Knowledge Discovery, Springer-Verlag, London, 1994.