A CLASSIFICATION STUDY OF ROUGH SETS GENERALIZATION

A Thesis
Submitted to the Faculty of Graduate Studies and Research
In Partial Fulfillment of the Requirements
for the Degree of
Master of Science
in
Computer Science
Faculty of Art and Science
Lakehead University

By
Han Yin Shi
Thunder Bay, Ontario
April 2004

Your file Votre référence
ISBN: 0-612-96987-8
Our file Notre référence
ISBN: 0-612-96987-8

Table of Contents

List of Figures

List of Tables

Abstract

In the development of rough set theory, many different interpretations and formulations have been proposed and studied. One can classify the studies of rough sets into algebraic and constructive approaches. While algebraic studies focus on the axiomatization of rough set algebras, the constructive studies concern with the construction of rough set algebras from other well known mathematical concepts and structures. The constructive approaches are particularly useful in the real applications of rough set theory. The main objective of this thesis to provide a systematic review existing works on constructive approaches and to present some additional results. Both constructive and algebraic approaches are first discussed with respect to the classical rough set model. In particular, three equivalent constructive definitions of rough set approximation operators are examined. They are the element based, the equivalence class based, and the subsystem based definitions. Based on the element based and subsystem based definitions, generalized rough set models are reviewed and summarized. One can extend the element based definition by using any binary relations instead of equivalence relations in the classical rough set model. Many classes of rough set models can be established based on the properties of binary relations. The subsystem based definition can be extended in the set-theoretical setting, which leads to rough set models based on Pawlak approximation space, topological space, and closure system. Finally, the connections between the algebraic studies, relation based, and subsystem based formulations are established.

Acknowledgement

I would like to express my gratitude to my thesis co-supervisors, Professor Yiyu Yao and Professor Ruizhong Wei, who led me into the research of rough set theory and guide me through the whole period of master study. Their knowledge, inspiration, energy and patience will always serve to me as an example of good researcher.

I would like to thank my external examiner Professor Yanqing Zhang. In addition, I would like to extend my thanks to many professors in Computer Science and Mathematics Departments. I deeply cherish my experience in Lakehead University during the period of master study. I would like to thank the Office of Graduate Studies, NSERC for financial support. Last but not the least, I am deeply grateful to my family for their endless love, concern and support.

Chapter 1

INTRODUCTION

The theory of rough sets was proposed by Pawlak more than 20 years ago [25]. There is a rapid grow of interest in the theory and its applications [14, 20, 23, 28]. Many interpretations and formulations of the theory have been proposed and studied [2, 48, 40, 43, 46, 53, 55]. Each of them captures a particular aspect of the theory. The co-existence of the many views and formulations shows the flexibility and wide applicability of the theory. This chapter reviews the basic concepts of the classical rough set model, i.e., the Pawlak rough set model. The formulation of rough set theory is divided into constructive and algebraic methods [42]. For constructive methods, there exist at least three definitions of rough set approximation operators [43, 46]. The algebraic methods focus on axioms of rough approximation operators [19, 48]. The relationships between the constructive and algebraic approaches can be established [48, 50, 53]. It should be commented that for an easy comparison with other mathematical systems and related theories, in this thesis a same property is often labeled differently in different contexts.

1.1 Pawlak Approximation Spaces

Let U denote a finite and non-empty set called the universe. Suppose $E \subseteq U \times U$ is an equivalence relation on U. That is, E satisfies the following three conditions:

(i) reflexibility : $\forall x \in U(xEx)$,

(ii) symmetry : $\forall x, y \in U(xEy \Longrightarrow yEx)$,

(iii) transitivity : $\forall x, y, z \in U(xEy \wedge yEz \Longrightarrow xEz)$.

The pair $apr = (U, E)$ is called a Pawlak approximation space. In an approximation space $apr = (U, E)$, the equivalence relation E partitions the set U into disjoint subsets. Such a partition of the universe is denoted by U/E. For an element $x \in U$, the equivalence class containing x is defined by:

$$[x]_E = \{y \mid xEy\}. \tag{1.1}$$

The equivalence relation reflects the relationships between elements in the universe U. It captures the indiscernibility of objects defined by available information [25, 27]. If two elements x, y in U belong to the same equivalence class, we say that x and y are indistinguishable. In real applications, the equivalence relation can be defined in an information table [27, 52]. The equivalence classes of E are called the elementary sets in the approximation space $apr = (U, E)$. The empty set $\emptyset$ and the union of one or more elementary sets is called a composed set [25, 27]. The family of all composed sets is denoted by $\mathrm{Com}(U)$, which forms a Boolean algebra. In general, $\mathrm{Com}(U) \neq 2^U$, where 2^U is the power set of U. That is, some subsets of U are not composed sets.

1.2 Rough Set Approximations

An important implication of the equivalence relation is the granulation of the universe. Under the equivalence relation, we can not differentiate two elements x

and y if xEy. That is, the available information or knowledge forces us to think the equivalence class $[x]_E$ as a whole, instead of many individuals. This implies that we can only describe a subset of elements which is an equivalence class of E, and, more generally, the subsets in $\text{Com}(U)$. For this reason, we also refer to a subset in $\text{Com}(U)$ as a definable, measurable, or interpretable set, depending on the particular context of applications [46]. Recall that in general, we have $\text{Com}(U) \neq 2^U$. This suggests that we can not have a precise characterization of a subset X not in $\text{Com}(U)$. The question of how to describe X roughly or approximately is to the development of the rough set theory. Institutively, this question can be answered as follows. Given a subset $X \subseteq U$ such that X is not a composed set, one can find a family of sets in $\text{Com}(U)$ that are subsets of X, and one can also find another family of sets in $\text{Com}(U)$ that are supersets of X. Hence, one can infer information about X based on the two families of sets in $\text{Com}(U)$, because X falls within the sets in two families. We observe that the first family has a maximum element and the second has a minimum element. Those two extreme elements are used to define a pair of approximations called the lower and upper approximation of X. More specifically, they are defined by [25]:

$$\begin{aligned} \underline{apr}(X) &= \text{the largest subset of } X \text{ which is an element of } \text{Com}(U), \\ \overline{apr}(X) &= \text{the smallers superset of } X \text{ which is an element of } \text{Com}(U). \end{aligned} \quad (1.2)$$

This definition is well defined and is applicable to any subset of U. In the case when $X \in \text{Com}(U)$, the lower and upper approximations are in fact X itself. This agrees with the interpretation of $\text{Com}(U)$. In the next section, we further examine the rough set approximations based on three constructive definitions.

1.3 Constructive Studies of Rough Set Theory

We examine three equivalent definitions that offer different interpretations of rough set approximations. They can be used to establish connections between rough set theory and other related theories, such as modal logic, topological operators, belief functions, and so on [37, 48, 49, 50, 53]. Furthermore, in the following two chapters, we will show that they allow derive different generalizations of the classical rough set theory. More specifically, the element based definition enables us to generalize the theory by using an arbitrary binary relation, the subsystem based definition enables us to generalize the theory using other mathematical structures such as topological spaces, closure systems, Boolean algebra, Lattices, and Posets [43, 46, 50].

1.3.1 Subsystem based definition

The subsystem based defined was first used by Pawlak in the study of the topological characteristics of rough set approximations [25]. In an approximation space $apr = (U, E)$, the set of composed sets $\mathrm{Com}(U)$ is in fact an σ-algebra with U/E as its basis. That is, $\mathrm{Com}(U)$ is closed with respect to set complement, intersection and union. Thus, we also denote this σ-algebra as $\sigma(U/E)$. The rough set approximations can be immediately defined by [25]:

$$\begin{aligned} \underline{apr}(X) &= \bigcup\{Y \mid Y \in \sigma(U/E), Y \subseteq X\}, \\ \overline{apr}(X) &= \bigcap\{Y \mid Y \in \sigma(U/E), Y \supseteq X\}, \end{aligned} \tag{1.3}$$

According to this definition, $\underline{apr}(X)$ is the largest definable set in the subsystem $\sigma(U/E)$ that is contained in X, $\overline{apr}(X)$ is the smallest definable subset in $\sigma(U/E)$ that contains the set X. A Pawlak approximation space defines uniquely a topological space $(U, \mathrm{Com}(U))$, in which $\mathrm{Com}(U)$ is the family of all open and closed sets [25]. Thus, the lower approximation is related to the interior, and the upper approximation is related to the closure, of a set in the topological space $U, \mathrm{Com}(U))$. This proves a topological interpretation of rough set approximations [25].

1.3.2 Element based definition

With respect to the lower and upper approximations, Pawlak introduced two types of memberships [25]. If $x \in \underline{apr}(X)$, we say that X surely belongs to X in apr and hence x is a strong member. If $x \in \overline{apr}(X)$, we say that x possibly belongs to X in apr, and hence x is a weak member [25]. The element based definition focuses on the conditions for strong and weak members. For an arbitrary set $X \subseteq U$, its lower and upper approximation is defined by:

$$
\begin{aligned}
\underline{apr}(X) &= \{x \in U \mid [x]_E \subseteq X\} \\
&= \{x \in U \mid \text{for all } y \in U, xEy \text{ implies } y \in X\} \\
&= \{x \in U \mid \forall y[y \in [x]_E \Longrightarrow y \in X]\}, \\
\overline{apr}(X) &= \{x \in U \mid [x]_E \cap X \neq \emptyset\} \\
&= \{x \in U \mid \text{there exists a } y \in U \text{ such that } xEy \text{ and } y \in X\} \\
&= \{x \in U \mid \exists y[y \in [x]_E, y \in X]\}. \qquad (1.4)
\end{aligned}
$$

An element $x \in U$ belongs to the lower approximation of X if all its equivalent elements belong to X. It belongs to the upper approximation of X if at least one of its equivalent elements belongs to X. The element based definition relates rough set approximation to the necessity and possibility in modal logics [50, 48].

1.3.3 Equivalence class based definition

In an approximation space, the equivalence classes may be considered as the smallest definable subsets. All composed set can be expressed as a union of equivalence classes. Thus, equivalence classes are the basic building blocks for rough set approximations. Being composed sets, the lower and upper approximations of a set can be expressed as unions of some equivalence classes. This offers the following equivalence class based definition:

$$\begin{aligned} \underline{apr}(X) &= \bigcup\{[x]_E \mid [x]_E \subseteq X\}, \\ \overline{apr}(X) &= \bigcup\{[x]_E \mid [x]_E \cap X \neq \emptyset\}. \end{aligned} \tag{1.5}$$

That is, $\underline{apr}(X)$ is the union of equivalence classes which are subsets of X, $\overline{apr}(X)$ is the union of equivalence classes which have a nonempty intersection with X. The equivalence class based definition can be used generalize rough set theory based coverings of a universe [47, 46, 55].

1.4 Properties of Rough Set Approximations

The three constructive definitions give the same rough set approximations, but different semantic interpretations. This section summarizes the main properties of such approximations. For any subsets $X, Y \subseteq U$, the lower approximation $\underline{apr}$ satisfies properties:

(L0) $\underline{apr}(X) = \sim\overline{apr}(\sim X)$,

(L1) $\underline{apr}(U) = U$,

(L2) $\underline{apr}(X \cap Y) = \underline{apr}(X) \cap \underline{apr}(Y)$,

(L3) $\underline{apr}(X \cup Y) \supseteq \underline{apr}(X) \cup \underline{apr}(Y)$,

(L4) $X \subseteq Y \Longrightarrow \underline{apr}(X) \subseteq \underline{apr}(Y)$,

(L5) $\underline{apr}(\emptyset) = \emptyset$,

(L6) $\underline{apr}(X) \subseteq X$,

$$\text{(L7)} \quad X \subseteq \underline{apr}(\overline{apr}(X)),$$

$$\text{(L8)} \quad \underline{apr}(X) \subseteq \underline{apr}(\underline{apr}(X)),$$

$$\text{(L9)} \quad \overline{apr}(X) \subseteq \underline{apr}(\overline{apr}(X)),$$

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

$$\text{(U0)} \quad \overline{apr}(X) = \ \sim\underline{apr}(\sim X),$$

$$\text{(U1)} \quad \overline{apr}(\emptyset) = \emptyset,$$

$$\text{(U2)} \quad \overline{apr}(X \cup Y) = \overline{apr}(X) \cup \overline{apr}(Y),$$

$$\text{(U3)} \quad \overline{apr}(X \cap Y) \subseteq \overline{apr}(X) \cap \overline{apr}(Y),$$

$$\text{(U4)} \quad X \subseteq Y \Longrightarrow \overline{apr}(X) \subseteq \overline{apr}(Y),$$

$$\text{(U5)} \quad \overline{apr}(U) = U,$$

$$\text{(U6)} \quad X \subseteq \overline{apr}(X),$$

$$\text{(U7)} \quad \overline{apr}(\underline{apr}(X)) \subseteq X,$$

$$\text{(U8)} \quad \overline{apr}(\overline{apr}(X)) \subseteq \overline{apr}(X),$$

$$\text{(U9)} \quad \overline{apr}(\underline{apr}(X)) \subseteq \underline{apr}(X),$$

where $\sim X = U - X$ denotes the set complement of X. Moreover, lower and upper approximations obey properties:

$$\text{(K)} \quad \underline{apr}(\sim X \cup Y) \subseteq \ \sim\underline{apr}(X) \cup \underline{apr}(Y),$$

$$\text{(K}'\text{)} \quad \sim\overline{apr}(X) \cap \overline{apr}(Y) \subseteq \ \overline{apr}(\sim X \cap Y),$$

$$\text{(D)} \quad \underline{apr}(X) \subseteq \overline{apr}(X).$$

Properties (L0) and (U0) state that two approximations are dual to each other. Hence, properties with the same number may be regarded as dual properties. These properties are not independent. For example, property (L2) implies property (L3) and (U2) implies property (U3). Properties (L8), (L9), (U8) and (U9) are expressed in terms of set inclusion. The standard version using set equality can be derived from (L0)-(L9) and (U0)-(U9). For example, it follows from (L6) and (L8) that $\underline{apr}(X) = \underline{apr}(\underline{apr}(X))$.

1.5 Algebraic Studies of Rough Set Theory

One may interpret $\underline{apr}, \overline{apr} : 2^U \longrightarrow 2^U$ as a pair of dual unary set-theoretic operators called approximation operators. The system $(2^U, \sim, \underline{apr}, \overline{apr}, \cap, \cup)$ is called a rough set algebra [48]. It is an extension of the set algebra $(2^U, \sim, \cap, \cup)$ with added operators. By interpreting rough set approximations as unary operators, one can have algebraic studies of the rough set theory. Instead of of constructing rough set operators first, and then studying their properties, one can study directly the rough set algebras by imposing on properties given in the last section. That is, we define a pair of dual approximation operators and states properties that must be satisfied by the operators. In an algebraic method, we focus on an algebraic system $(2^U, \cap, \cup, \sim, \mathbf{L}, \mathbf{H})$, where $(2^U, \cap, \cup, \sim)$ is the set algebra, and $\mathbf{L}, \mathbf{H} : 2^U \longrightarrow 2^U$ are two unary operators on the power set 2^U. A pair of unary set-theoretic operators $\mathbf{L}, \mathbf{H}$ are called dual operators, if they satisfy the properties:

$$\text{(L0)} \qquad \mathbf{L}(X) = \sim\mathbf{H}(\sim X),$$

$$\text{(H0)} \qquad \mathbf{H}(X) = \sim\mathbf{L}(\sim X),$$

With respect to the properties given in the last section, we consider the following list of axioms of approximation operators L and H:

$$\text{(K)} \quad \mathbf{L}(\sim X \cup Y) \subseteq \sim \mathbf{L}(X) \cup \mathbf{L}(Y),$$

$$\text{(K}'\text{)} \quad \sim \mathbf{H}(X \cap \mathbf{H}Y) \subseteq \mathbf{H}(\sim X \cap Y),$$

$$\text{(L1)} \quad \mathbf{L}(U) = U,$$

$$\text{(L2)} \quad \mathbf{L}(X \cap Y) = \mathbf{L}(X) \cap \mathbf{L}(Y),$$

$$\text{(H1)} \quad \mathbf{H}(\emptyset) = \emptyset,$$

$$\text{(H2)} \quad \mathbf{H}(X \cup Y) = \mathbf{H}(X) \cup \mathbf{H}(Y),$$

$$\text{(D)} \quad \mathbf{L}(X) \subseteq \mathbf{H}(X);$$

$$\text{(T)} \quad \mathbf{L}(X) \subseteq (X),$$

$$\text{(T}'\text{)} \quad X \subseteq \mathbf{H}(X);$$

$$\text{(B)} \quad X \subseteq \mathbf{LH}(X),$$

$$\text{(B}'\text{)} \quad \mathbf{HL}(X) \subseteq X;$$

$$\text{(4)} \quad \mathbf{L}(X) \subseteq \mathbf{LL}(X),$$

$$\text{(4}'\text{)} \quad \mathbf{HH}(X) \subseteq \mathbf{H}(X);$$

$$\text{(5)} \quad \mathbf{H}(X) \subseteq \mathbf{LH}(X),$$

$$\text{(5}'\text{)} \quad \mathbf{HL}(X) \subseteq \mathbf{L}(X).$$

The axioms in the list are relabeled by following the convenient in modal logic [4, 48, 50]. Based on those axioms, we can study and classify different rough set algebras [42].

1.6 Connections of Constructive and Algebraic Studies

The connection between constructive and algebraic methods can be established by finding a set of axioms so that there exists an equivalence relation that produces the same approximation operators [19]. The following theorem gives a set of such axioms [48].

Theorem 1.1 *Suppose* $\mathbf{L}, \mathbf{H} : 2^U \longrightarrow 2^U$ *is a pair of dual operators. If* $\mathbf{H}$ *satisfies axioms* (H1), (H2), (T′), (4′) *and* (B), *then there exists an equivalence relation* E *on* U *such that for all* $X \subseteq U$, $L(X) = \underline{apr}(X)$ *and* $H(A) = \overline{apr}(X)$, *where* $\underline{apr}$ *and* $\overline{apr}$ *are the approximation operators defined by the equivalence relation* E.

This theorem can be extended when studying other rough set algebras [48, 50]. Some authors also studied operators **L** and **H** that are not dual to each other [29, 30, 55].

1.7 Organization of the Thesis

The main contribution is a more complete and coherent study of constructive approaches of rough set theory. Our emphasis is on putting the majority of existing studies in a unified framework. This enables us to gain more insights into existing studies, and at the same time, to obtain new results that enriches the constructive approaches. In order to achieve such objectives, the thesis is organized as follows. In Chapter 1, by reviewing the existing studies on constructive and algebraic approaches, we establish a unified framework in which one can construct generalized rough set models. The generalized models can be obtained by extending the three constructive definitions. In Chapter 2, we briefly review and summarize constructive approaches based on the use of an arbitrary binary relation, instead of an equivalence relation. For such generalizations, element based defining is used. A connection between properties binary relations and rough set approximation operators can be established. Different rough set models can be derived. In Chapter 3, we examine generalized rough set models constructed by using the subsystem based definition. The discussion is developed in the set-theoretic setting. In particular, Pawlak topology, topological spaces and closure systems are used. The generalization of rough set theory using other mathematical structures greatly enriches the theory. In Chapter 4, we study the connections between algebraic studies, binary relation based and subsystem based formulations. The results of Chapters 2 and 3 are thus linked together. In Chapter 5, we summarize this thesis and point out some further research topics.

Chapter 2

CONSTRUCTION OF GENERALIZED ROUGH SETS BASED ON BINARY RELATIONS

The element based definition can be easily extended to construct generalized rough set models based on an arbitrary binary relations [39, 50]. As shown by Yao and Lin [50], such a study can be carried out by drawing results from modal logic [4].

2.1 Generalized Approximation Spaces

The Pawlak approximation space is defined by an equivalence relation. This unnecessarily limits the flexibility and applicability of the theory. In many real applications, the requirement of the transitivity may be too strong [33, 34]. Hence, it is necessary to consider non-equivalence relations. Let $R \subseteq U \times U$ be a binary relation on U. The pair $apr = (U, R)$ is called a generalized approximation space. Depending on the properties of the binary relation R, one can derived many classes of approximation spaces. For two elements $x, y \in U$, if xRy, we say that y is R-related to x. The physical meaning is that x and y are somewhat semantically related. It can be interpreted as indistinguishable, accessibility, similarity, or connectivity, depending on the context. A binary relation can be equivalently defined by a mapping from U to the power set 2^U:

$$xR = \{y \mid xRy\}.$$

The set $xR \subseteq U$ consists of elements of U that are R-related to x. It is called the successor neighborhood of x. In general, one may also other types of neighborhoods []. When an equivalence relation E is used, the successor neighborhood xE is in fact the equivalence class $[x]_E$ containing x.

2.2 Generalized Approximation Operators

Many generalized models have been developed for the theory of rough set by constructing different types of approximation operators based on non-equivalence relations [39, 50]. As shown by Yao [47, 46], by substituting $[x]_E$ with xR, one can obtain generalized approximation operators by extending either the element based definition or the equivalence class based definition. We consider the extension of element based definition. By replacing equivalence classes with successor neighborhoods in the element based definition, we define generalized approximation operators as:

$$\begin{aligned} \underline{apr}(X) &= \{x \mid \forall y[xRy \Longrightarrow y \in X]\} \\ &= \{x \mid xR \subseteq X\}, \\ \overline{apr}(X) &= \{x \mid \exists y[xRy, y \in X]\}. \\ &= \{x \mid xR \cap X \neq \emptyset\}. \end{aligned} \tag{2.6}$$

The set $\underline{apr}(X)$ consists of those elements whose R-related elements are all in X, and $\overline{apr}(X)$ consists of those elements such that at least one of whose R-related elements is in X. The generalized approximation operators are defined in a similar manner as that of the necessity and possibility operators in modal logic [50]. It follows that we can employ the results from modal logic to study generalized rough set models. As a first step, we adopt the same labeling system from Chellas [4] for naming approximation operators:

$$
\begin{array}{ll}
\text{(K)} & \underline{apr}(\sim X \cup Y) \subseteq \ \sim \underline{apr}(X) \cup \underline{apr}(Y), \\
\text{(D)} & \underline{apr}(X) \subseteq \overline{apr}(X), \\
\text{(T)} & \underline{apr}(X) \subseteq X, \\
\text{(T}') & X \subseteq \overline{apr}(X); \\
\text{(B)} & X \subseteq \underline{apr}(\overline{apr}(X)), \\
\text{(B}') & \overline{apr}(\underline{apr}(X)) \subseteq X; \\
\text{(4)} & \underline{apr}(X) \subseteq \underline{apr}(\underline{apr}(X)), \\
\text{(4}') & \overline{apr}(\overline{apr}(X)) \subseteq \overline{apr}(X); \\
\text{(5)} & \overline{apr}(X) \subseteq \underline{apr}(\overline{apr}(X)); \\
\text{(5}') & \overline{apr}(\underline{apr}(X)) \subseteq \underline{apr}(X).
\end{array}
$$

The Pawlak approximation operators satisfy all those properties. Generalized approximation operators do not necessarily satisfy them. Property (K) does not depend on any particular binary relation. The other properties depends on the properties of binary relations. Each of the properties (D)-(5) corresponds to a property of the binary relation. Table 2.1 summarizes the connections between properties of binary relations and properties of approximation operators [43, 47]. More specifically, the properties (D), (T), (B), (4), and (5) correspond to serial, reflexive, symmetric, transitive, and Euclidean binary relations, respectively. By combining these properties, one can construct many distinct rough set models. The respective models are named

Property of binary relation		Property of approximation operators
none		(K)
serial:	for all $x \in U$, there exists a $y \in U$, such that $y \in xR$	(D)
reflexive:	for all $x \in U$, $x \in xR$	(T)
symmetric:	for all $x, y \in U$, $x \in yR \Longrightarrow y \in xR$	(B)
transitive:	for all $x, y, z \in U$, $[y \in xR, z \in yR] \Longrightarrow z \in xR$	(4)
Euclidean:	for all $x, y, z \in U$, $[y \in xR, z \in xR] \Longrightarrow y \in zR or z \in yR$	(5)

Table 2.1: Relationships between properties of binary relations and approximation operators

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 the binary relation is serial, one obtains a rough set algebra that is an interval structure [38]. If the binary relation is reflexive and transitive, one obtains a topological rough set algebra [16, 48]. If R is reflexive and symmetric, i.e., R is a compatibility relation, properties (K), (D), (T) and (B) hold. This model is labeled by KTB. Property (D) does not explicitly appear in the label because it is implied by (T). Pawlak rough set model is labeled by KT5, which is also commonly known as S5 in modal logic. It should be noted that two different subsets of properties in Table 2.1 may produce the same rough set model. By considering all subsets of the property set, {none, serial, reflexive, symmetric, transitive, Euclidean}, one can derived at least fifteen distinct classes of rough set models [50]. Figure 2.1, adopted from Chellas [4] and Marchal [21], 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

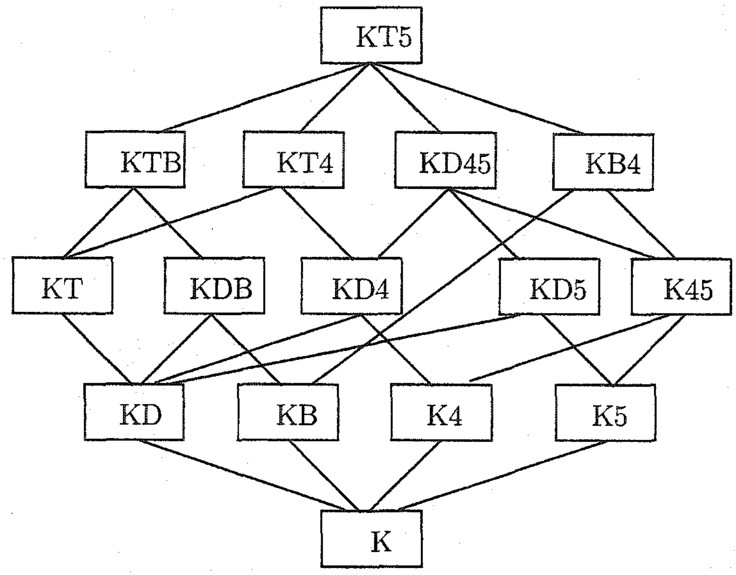

Figure 2.1: Rough set models

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.

Chapter 3

CONSTRUCTION OF GENERALIZED ROUGH SET MODELS BASED ON SUBSYSTEMS

In this chapter, we review and discuss generalizations of rough set theory based on subsystem based definition in the set-theoretical setting. The subsystem of the power set of the universe, such as close-open topological spaces, topological spaces, and closure systems, are used.

3.1 Pawlak Topological Space

The Pawlak approximation space, $apr = (U, E)$, is characterized by an equivalence relation $E \subseteq U \times U$ defined on the finite and non-empty universe U. The equivalence relation E partitions the universe U into pairwise disjoint subsets called the partition of the universe, written U/E. The equivalence classes and the empty set $\emptyset$ are called elementary sets. The empty set and the unions of elementary sets are called composed sets, or the definable, observable, measurable sets [25, 48]. The family of all definable sets is denoted by $\mathrm{Com}(U)$. The system $\mathrm{Com}(U)$ is closed under set complement, intersection and union. It is a subsystem of the power set 2^U. In fact, it is an σ-algebras of subsets of U. There is a closely connection between Pawlak approximation spaces and special type of topological spaces where the family of open and closed sets are the same. For a Pawlak approximation space on a finite universe, we have a topological space $(U, \text{Com}(U)$. Moreover, the family of clopen sets, i.e., closed and open sets, the same as the family of closed sets, and the family of the open sets are the same, namely, they all equal to $\text{Com}(U)$. This type of topological space is sometimes called Pawlak topology [18]. For a finite set U, there is a natural correspondence between the set of all equivalence relations on U and the set of topologies on U in which each closed set is open (i.e., Pawlak topology) [9]. We can restate the subsystem based definition as follows:

(def1)
$$\underline{apr}(X) = \bigcup\{Y \mid Y \in \text{Com}(U), Y \subseteq X\},$$
$$\underline{apr}(X) = \bigcap\{Y \mid Y \in \text{Com}(U), X \subseteq Y\}.$$

The greatest clopen set contained in X is the lower approximation of X and the least clopen set containing X is the upper approximation of X. The pair of lower and upper approximations may be interpreted as two unary set-theoretic operators $U : \underline{apr}, \overline{apr} : 2^U \longrightarrow 2^U$. Under this view, the rough set theory can be regarded as extension of set theory with two additional unary set-theoretic operators $\underline{apr}(X)$ and $\overline{apr}(X)$. Together with the standard set-theoretic operators negative $\sim$, intersection $\cap$ and union $\cup$, the system $(2^U, \sim, \underline{apr}, \overline{apr}, \cap, \cup)$ is called a Pawlak rough set algebra. The approximation operators are in fact the interior and closure operators in the Pawlak topology $(U, \text{Com}(U)$. The set of fixed points of the two operators are defined by:

$$\begin{aligned} \mathrm{O}(U) &= \{X \mid \underline{apr}(X) = X\} \\ &= \{X \mid \overline{apr}(X) = X\} \\ &= \mathrm{C}(U) \\ &= \mathrm{Com}(U). \end{aligned} \tag{3.7}$$

That is, we can recover the set of clopen sets in a Pawlak topology by using the fixed points of approximation operators.

3.2 Topological Space

In a Pawlak topology, the family of all open and closed sets [25]. This is not necessarily true for an arbitrary topological space. The subsystem based definition of approximation operators can be extended for any topological space [5, 16, 36, 48]. Let $(U, \mathrm{O}(U))$ be a topological space, where $\mathrm{O}(U) \subseteq 2^U$ is a family of subsets of U satisfying the following axioms:

(O1) $\emptyset \in \mathrm{O}(U), U \in \mathrm{O}(U)$;

(O2) $\mathrm{O}(U)$ is closed under union, i.e.,

for any subsystem $\mathcal{D} \subseteq \mathrm{O}(U)$, we have $\bigcup \mathcal{D} \in \mathrm{O}(U)$;

(O3) $\mathrm{O}(U)$ is closed under finite intersection, i.e.,

for any $A, B \in \mathrm{O}(U)$, we have $A \cap B \in \mathrm{O}(U)$.

Elements of $\mathrm{O}(U)$ are called open sets. A set in U is called closed set if and only if its complement set is a open set. This establishes the duality between the closed sets and open sets by means of set complement. One can easily deduce the properties of closed sets. The family of all closed sets $C(U) = \{\sim X \mid X \in O(U)\}$ is characterized by the following axioms:

(C1) $\emptyset \in C(U), U \in C(U)$;

(C2) $C(U)$ is closed under intersection, i.e.,

for any subsystem $\mathcal{D} \subseteq C(U)$, we have $\bigcap \mathcal{D} \in C(U)$;

(C3) $C(U)$ is closed under finite union, i.e.,

for any $A, B \in C(U)$, we have $A \cup B \in C(U)$.

In general, the family of open sets is different from the family of closed sets, that is $O(U) \neq C(U)$. We denote

$$CO(U) = O(U) \cap C(U) \quad (3.8)$$

the set of all clopen (simultaneously closed and open) sets. For a finite universe U, axioms (O2) and (C2) may be simply stated in terms of finite union and intersection. The approximation operators can be constructed based on the subsystem based definition [3, 46]. The system $O(U)$ can be viewed as the sets of inner definable sets, and the system $O(U)$ as the set of outer definable elements. By extending definition (**defl**), a subset $X \subseteq U$ can be approximated from below by an element of $O(U)$, and from above by an element of $C(U)$:

(**defla**)
$$\underline{apr}(X) = \bigcup\{Y \mid Y \in O(U), Y \subseteq X\},$$
$$\overline{apr}(X) = \bigcap\{Y \mid Y \in C(U), X \subseteq Y\}.$$

That is, $\underline{apr}(X)$ the largest open set contained in X, and $\overline{apr}(X)$ is the smallest closed set containing X. They in fact define a pair of Kuratowski topological interior and closure operators characterized by the axioms:

(i1) $\underline{apr}(X \cap Y) = \underline{apr}(X) \cap \underline{apr}(Y)$, (i2) $\underline{apr}(X) \subseteq X$,

(i3) $\underline{apr}(\underline{apr}(X)) = \underline{apr}(X)$,

(i4) $\underline{apr}(U) = U$,

and

(c1) $\overline{apr}(X \cup Y) = \overline{apr}(X) \cup \overline{apr}(Y)$,

(c2) $X \subseteq \overline{apr}(X)$,

(c3) $\overline{apr}(\overline{apr}(X)) = \overline{apr}(X)$,

(c4) $\overline{apr}(\emptyset) = \emptyset$.

Conversely, given a pair of dual approximation operators, $\underline{apr}, \overline{apr} : 2^U \longrightarrow 2^U$, satisfying axioms (i1)-(i4) and axioms (c1)-(c4), respectively, the sets of their fixed points:

$$\begin{aligned} \mathrm{O}(U) &= \{X \mid \underline{apr}(X) = X\}, \\ \mathrm{C}(U) &= \{X \mid \overline{apr}(X) = X\}, \end{aligned} \tag{3.9}$$

are families of open and closed sets of a topological space.

3.3 Closure System

The notion of closed sets in a topological space may be further generalized by removing some of axioms from (c1)-(c4). A family $\mathcal{C}(U)$ of subsets of U is called a closure system if $U \in \mathcal{C}(U)$ and is closed under intersection [6]. That is,

(C1′) $U \in \mathcal{C}(U)$;

(C2) $\mathcal{C}(U)$ is closed under intersection, i.e.,

for any subsystem $\mathcal{D} \subseteq \mathcal{C}(U)$, we have $\bigcap \mathcal{D} \in \mathcal{C}(U)$.

By collecting the complements of members of $\mathcal{C}(U)$, we obtain another system $\mathcal{O}(U) = \{\neg X \mid X \in \mathcal{C}(U)\}$. According to properties of $\mathcal{C}(U)$, the system $\mathcal{O}(U)$ contains the empty set $\emptyset$ and is closed under union, namely,

(O1′) $\quad U \in \mathcal{O}(U)$;

(O2) $\quad \mathcal{O}(U)$ is closed under union, i.e.,

for any subsystem $\mathcal{D} \subseteq \mathcal{O}(U)$, we have $\bigcup \mathcal{D} \in \mathcal{O}(U)$.

The pair of systems $\mathcal{O}(U)$ and $\mathcal{O}(U)$ correspond to the families of open sets and closed sets in a topological space. Definition (**def1**) may be generalized to produce approximation operators in a closure system:

$$\textbf{(def1b)} \qquad \underline{apr}(X) = \bigcup\{Y \mid Y \in \mathcal{O}(U), Y \subseteq X\},$$

$$\underline{apr}(X) = \bigcap\{Y \mid Y \in \mathcal{C}(U), X \subseteq Y\}.$$

The upper approximation operator is in fact a closure operator satisfying the following axioms:

(j1) $\quad$ If $X \subseteq Y$, then $\overline{apr}(X) \subseteq \overline{apr}(Y)$,

(j2) $\quad X \subseteq \overline{apr}(X)$,

(j3) $\quad \overline{apr}(\overline{apr}(X)) = \overline{apr}(X)$.

The lower approximation operator satisfies the properties:

(j1′) $\quad$ If $X \subseteq Y$, then $\underline{apr}(X) \subseteq \underline{apr}(Y)$,

(i2′) $\quad \underline{apr}(X) \subseteq X$,

(j3′) $\quad \underline{apr}(\underline{apr}(X)) = \underline{apr}(X)$.

Properties (j2) and (j3) are in fact part of the properties of closure operator and (j2′) and (j3′) are part of the properties of interior operator in a topological space.

Rough set approximation operators based on closure system are weaker those defined in a topological space. Conversely, for a closure operator $\overline{apr} : 2^U \longrightarrow 2^U$ satisfying axioms (j1), (j2), and (j3), the set of fixed points of $\overline{apr}$:

$$\mathcal{C}(U) = \{X \mid \overline{apr}(X) = X\}, \tag{3.10}$$

is a closure system. Similar results can be stated between the system $\mathcal{O}(U)$:

$$\mathcal{O}(U) = \{X \mid \underline{apr}(X) = X\}, \tag{3.11}$$

and the dual operator $\underline{apr}(X)$.

3.4 Boolean Algebra and Lattice

In the Pawlak topological space, the power set 2^W is a Boolean algebra, and the family of composed set $\mathrm{Com}(U)$ is a sub-Boolean algebra. Thus, the subsystem based definition can be easily generalized using Boolean algebra and lattice [45, 46, 53]. Suppose $(\mathcal{B}, \neg, \wedge, \vee, 0, 1)$ is a finite Boolean algebra and $(\mathcal{B}_0, \neg, \wedge, \vee, 0, 1)$ is a sub-Boolean algebra. One may approximate an element of $\mathcal{B}$ using elements of $\mathcal{B}_0$:

$$\textbf{(Ldef1)} \qquad \begin{aligned} i(x) &= \bigvee \{y \mid y \in \mathcal{B}_0, y \leq x\}, \\ c(x) &= \bigwedge \{y \mid y \in \mathcal{B}_0, x \leq y\}. \end{aligned}$$

Since any finite Boolean algebra is a complete Boolean algebra, thus the above definition is well defined. Moreover, operators i and c satisfy the following axioms:

(i1) $i(x \wedge y) = i(x) \wedge i(y)$,

(i2) $i(x) \leq x$,

(i3) $i(i(x)) = i(x)$,

(i4) $i(1) = 1$,

(i5) $c(x) = i(c(x))$,

and

$$
\begin{array}{ll}
\text{(c1)} & c(x \bigvee y) = c(x) \bigvee c(y), \\
\text{(c2)} & x \leq c(x), \\
\text{(c3)} & c(c(x)) = c(x), \\
\text{(c4)} & c(0) = 0, \\
\text{(c5)} & i(x) = c(i(x)),
\end{array}
$$

One may define a pair of approximation operators directly, and use their fixed points as inner and outer definable elements. The sets of fixed points of i and c are open and closed elements. The system $(\mathcal{B}_0, \neg, i, c, \bigwedge, \bigvee, 0, 1)$ is a topological Boolean algebra [32], which is an extension of Boolean algebra with added operators. Gehrke and Walker [10] considered a more generalized definition in which the Boolean algebra $\mathcal{B}$ is replaced by a completely distributive lattice. The subsystem $\mathcal{B}_0$ is a sub-lattice. A more detailed and systematic study of approximation operators in special types of lattice and posets, as well as examples, can be found in a recent paper by Cattaneo [2] and Yao [45]. A different formulation of approximation operators in poset can be found a paper by Iwinski [15].

Chapter 4

CONNECTIONS OF DIFFERENT FORMULATIONS

The main objective of this chapter is to give a synthesis of different formulations of rough set approximation operators. It is devoted to revealing interconnections between algebraic studies and the rough set approximation operators defined by binary relations and subsystems. There is a correspondence between an equivalence relation and a Pawlak topology, and between a reflexive, transitive relation and a topological space. On the other hand, a different type of correspondence can be obtained between a binary relation and a closure system.

4.1 Algebraic Studies and Binary Relation Based Formulation

Like the Pawlak rough set model, the constructive method can be related to the algebraic method by listing a set of axioms on the approximation operators for the existing of the respective binary relations. According to Table 2.1, each axiom corresponds to a property of the lower and upper approximation operators constructed from a binary relation having a particular property. Relationships between operators defined by axiomatic and constructive approaches are summarized below [43].

Theorem 4.2 *Suppose* $\mathbf{L}, \mathbf{H} : 2^U \longrightarrow 2^U$ *is a pair of dual approximation operators satisfy axioms* (L1) *and* (L2), *and axioms* (H1) *and* (H2). *There exists*

(a). *a serial relation* R on U,

(b). *a reflexive relation* R on U,

(c). *a symmetric relation* R on U,

(d). *a transitive relation* R on U,

(e). *an Euclidean relation* R on U,

such that $\mathbf{L}(X) = \underline{apr}_R(X)$ *and* $\mathbf{H}(X) = \overline{apr}_R(X)$ *for all* $X \subseteq U$, *if and only if* $\mathbf{L}$ *and* $\mathbf{H}$ *satisfy axiom:*

(a). (D);

(b). (T$'$);

(c). (B$'$);

(d). (4$'$);

(e). (5$'$);

where $\underline{apr}_R$ *and* $\overline{apr}_R$ *are the approximation operators defined by the binary relation* R.

In this theorem, the corresponding pair of condition and conclusion is linked together by the same letter. For instance, one can conclude that there exists a serial binary relation R such that $\mathbf{L}(X) = \underline{apr}(X)$ and $\mathbf{H}(X) = \overline{apr}(X)$ for all $X \subseteq U$, if and only if the pair of approximation operators $\mathbf{L}$ and $\mathbf{H}$ satisfies axiom (D). The theorem may view considered as a generalization of Theorem 1.1. The theorem establishes, in general, a link between approximation operators constructively defined and algebraically defined, respectively. Based on the one-to-one correspondence, we may use either the properties of binary relations or the axioms of operators to label particular classes of rough set algebras. The *only if* part of the theorem can be proved based on the discussion of Chapter 2. The *if* part can be proved by constructing a binary relation R based on the upper approximation operator $\mathbf{H}$ as follows [42, 48]:

$$xRy \Longleftrightarrow x \in \mathbf{H}(\{y\}). \tag{4.12}$$

That is, $xR = \{y \mid x \in \mathbf{H}(\{y\})\}$, and conversely $\mathbf{H}(\{y\}) = \{x \mid y \in xR\}$. Let

$$Ry = \{x \mid xRy\}, \tag{4.13}$$

denote the predecessor neighborhood of y. Then we have:

$$Ry = \mathbf{H}(\{y\}). \tag{4.14}$$

Given a upper approximation operator $\mathbf{H}$ satisfying a particular property, one can show that the constructed binary relation R satisfies the corresponding property.

4.2 Algebraic Studies and Subsystem Based Formulation

In the subsystem based constructive approaches, we define a pair of approximation operators using subsystems with certain properties. Conversely, given a pair of approximation operators, we can find a set of axioms that implies the existence of a pair of subsystems that produce the same operators. This would establish the connections between the algebraic studies and subsystem based formulation. Based on the discussion in Chapter 3 and results from topological spaces and closure systems, we can easily obtain the connections of algebraic studies and subsystem based formulation. The results are summarized as follows.

Theorem 4.3 *Suppose* $\mathbf{L}, \mathbf{H} : 2^U \longrightarrow 2^U$ *is a pair of dual approximation operators. There exists a Pawlak topology* $U, \mathrm{Com}(U))$ *such that* $\mathbf{L}(X) = \underline{apr}_P(X)$ *and* $\mathbf{H}(X) = \overline{apr}_P(X)$ *for all* $X \subseteq U$*, if and only if* $\mathbf{L}$ *and* $\mathbf{H}$ *satisfy following axioms:*

$$\begin{array}{ll} \text{(i1)} & \mathbf{L}(X \cap Y) = \mathbf{L}(X) \cap \mathbf{L}(Y), \\ \text{(i2)} & \mathbf{L}(X) \subseteq X, \\ \text{(i3)} & \mathbf{L}(\mathbf{L}(X)) = \mathbf{L}(X), \\ \text{(i4)} & \mathbf{L}(U) = U, \\ \text{(i5)} & \mathbf{H}(X) \subseteq \mathbf{L}(\mathbf{H}(X)), \end{array}$$

and

$$\begin{array}{ll} \text{(c1)} & \mathbf{H}(X \cup Y) = \mathbf{H}(X) \cup \mathbf{H}(Y), \\ \text{(c2)} & X \subseteq \mathbf{H}(X), \\ \text{(c3)} & \mathbf{H}(\mathbf{H}(X)) = \mathbf{H}(X), \\ \text{(c4)} & \mathbf{H}(\emptyset) = \emptyset, \\ \text{(c5)} & \mathbf{H}(\mathbf{L}(X)) \subseteq \mathbf{L}(X), \end{array}$$

where $\underline{apr}_P$ *and* $\overline{apr}_P$ *are the approximation operators defined by the subsystem based definition using the Pawlak topology* $(U, \mathrm{Com}(U))$.

The *only if* part follows the discussion of Chapter 3. The *if* part can be proved by first constructing the families of fixed points of the lower and upper approximation operators $\mathbf{L}$ and $\mathbf{H}$, and then showing the two families are the same. In fact, they are subsystem $\mathrm{Com}(U)$ of the Pawlak topology $U, \mathrm{Com}(U))$. The above theorem can be generalized to deal with subsystem based constructive methods using topological space and closure system.

Theorem 4.4 *Suppose* $\mathbf{L}, \mathbf{H} : 2^U \longrightarrow 2^U$ *is a pair of dual approximation operators. There exists a topology* $(U, \mathrm{O}(U))$ *such that* $\mathbf{L}(X) = \underline{apr}_T(X)$ *and* $\mathbf{H}(X) = \overline{apr}_T(X)$ *for all* $X \subseteq U$, *if and only if* $\mathbf{L}$ *and* $\mathbf{H}$ *satisfy axioms* (i1)-(i4) *and* (c1)-(c4), *where* $\underline{apr}_T$ *and* $\overline{apr}_T$ *are the approximation operators defined by the subsystem based definition using the topological space* $(U, \mathrm{O}(U))$.

Theorem 4.5 *Suppose* $\mathbf{L}, \mathbf{H} : 2^U \longrightarrow 2^U$ *is a pair of dual approximation operators. There exists a closure system* $(U, \mathcal{C}(U))$ *such that* $\mathbf{L}(X) = \underline{apr}_C(X)$ *and* $\mathbf{H}(X) = \overline{apr}_C(X)$ *for all* $X \subseteq U$, *if and only if* $\mathbf{L}$ *and* $\mathbf{H}$ *satisfy the following axioms:*

(j1) If $X \subseteq Y$, then $\mathbf{H}(X) \subseteq \mathbf{H}(Y)$,

(j2) $X \subseteq \mathbf{H}(X)$,

(j3) $\mathbf{H}(\mathbf{H}(X)) = \mathbf{H}(X)$,

and

(j1′) If $X \subseteq Y$, then $\mathbf{L}(X) \subseteq \mathbf{L}(Y)$,

(i2′) $\mathbf{L}(X) \subseteq X$,

(j3′) $\mathbf{L}(\mathbf{L}(X)) = \mathbf{L}(X)$,

where $\underline{apr}_C$ *and* $\overline{apr}_C$ *are the approximation operators defined by the subsystem based definition using the closure system* $(U, \mathcal{C}(U))$.

The proofs of those theorems easily follow from the results from topological spaces and closure systems [6]. The can be proved in the similar manner discussed earlier for Theorem 4.3.

4.3 Connection Between Relation Based and Subsystem Based Formulations

For a Pawlak rough set model with a finite universe, the family of clopen sets (i.e., closed and open sets) is defined by an equivalence relation, and vice versa [9]. Thus, the subsystem based definition and element based definition (i.e., the relation based definition) are the same. When an arbitrary binary is used, we need to seek similar type of connections. Suppose $R \subseteq U \times U$ is an arbitrary binary relation. The element based definition is given by [19, 48]:

$$
\begin{aligned}
\textbf{(def2a)} \qquad \underline{apr}_R(X) &= \{x \in U \mid xR \subseteq X\} \\
&= \{x \in U \mid \forall y[y \in xR \Longrightarrow y \in X]\}, \\
\overline{apr}_R(X) &= \{x \in U \mid xR \cap X \neq \emptyset\} \\
&= \{x \in U \mid \exists y[y \in xR, y \in X]\}.
\end{aligned}
$$

The subscript R indicates that the approximation operators are defined with respect to a binary relation R. Independent of the properties of R, $\underline{apr}_R$ satisfies axioms (i1) and (i4), while $\overline{apr}_R$ satisfies axioms (c1) and (c4). The sets of fixed points of the two operators are defined by:

$$
\begin{aligned}
\mathrm{O}_R(U) &= \{X \mid \underline{apr}_R(X) = X\}, \\
\mathrm{C}_R(U) &= \{X \mid \overline{apr}_R(X) = X\}.
\end{aligned} \tag{4.15}
$$

Obviously, $\mathrm{O}_R(U)$ contains U and satisfies axiom (O3), while $\mathrm{C}_R(U)$ contains $\emptyset$ and satisfies axiom (C3). Since $\mathrm{O}_R(U)$ is not closed under union and $\mathrm{C}_R(U)$ is not closed under intersection, we cannot use the generalized definition (**def1a**) with $\mathrm{O}_R(U)$ and $\mathrm{O}_R(U)$. In other words, for an arbitrary binary relation, the generalization (**def2a**) cannot be obtained by a generalization (**def1a**). Suppose T is a reflexive and transitive binary relation. The pair of approximation of operators given by definition (**def2a**) satisfies axioms (i2) and (i4), and axioms (c2) and (c4). In this case, their families of fixed points:

$$\begin{aligned} \mathrm{O}_T(U) &= \{X \mid \underline{apr}_T(X) = X\}, \\ \mathrm{C}_T(U) &= \{X \mid \overline{apr}_T(X) = X\}, \end{aligned} \tag{4.16}$$

are sets of open and closed sets of the topological space $(U, \mathrm{O}_T(U))$. Operators $\underline{apr}_T$ and $\overline{apr}_T$ are in fact the interior and closure operators defined using (**def1a**) with respect to the topological space $(U, \mathrm{O}_T(U)$. It can be easily verified that the successor neighborhood xT of $x \in U$ is an open set, namely, $\underline{apr}_T(xT) = xT$. By definition, we have $\underline{apr}_T(xT) = \{y \mid yT \subseteq xT\}$. Suppose $y \in \underline{apr}_T(xT)$. It follows from the definition that $yT \subseteq xT$. According to the reflexivity of T, it follows $y \in yT \subseteq xT$. Assume now that $y \in xT)$. By the transitivity of T we have $yT \subseteq xT$, that is, $y \in \underline{apr}_T(xT)$. Let

$$\begin{aligned} \mathrm{BO}_T(U) &= \{xT \mid x \in U\}, \\ \mathrm{BC}_T(U) &= \{\neg xT \mid x \in U\}. \end{aligned} \tag{4.17}$$

Any fixed point X of $\underline{apr}_T$, i.e., $\underline{apr}_T(X) = X \in \mathrm{O}_T(U)$, can be written in terms of a union of a subsystem of $\mathrm{BO}_T(U)$. More precisely, we have $X = \bigcup\{Y \mid Y \in \mathrm{BO}_T(U), Y \subseteq X\}$. The set $\mathrm{BO}_T(U)$ is a base of the topology $\mathrm{O}_T(U)$. With families $\mathrm{BO}_T(U)$ and $\mathrm{BC}_T(U)$, we obtain another definition of approximation operators:

$$\begin{aligned} \textbf{(def3a)} \qquad \underline{apr}_T(X) &= \bigcup\{Y \mid Y \in \mathrm{BO}_T(U), Y \subseteq X\} \\ &= \bigcup\{xT \mid x \in U, T_s(x) \subseteq X\} \\ &= \{x \in U \mid \exists y[x \in yT, yT \subseteq X]\}, \end{aligned}$$

$$
\begin{aligned}
\overline{apr}_T(X) &= \neg\{x \in U \mid \exists y[x \in yT, yT \subseteq \neg X]\} \\
&= \{x \in U \mid \forall y[x \in yT \Longrightarrow yT \not\subseteq \neg X]\} \\
&= \{x \in U \mid \forall y[x \in rT \Longrightarrow yT \cap X \neq \emptyset]\} \\
&= \bigcap\{\neg xT \mid x \in U, X \subseteq \neg xT\} \\
&= \bigcap\{Y \mid Y \in \mathrm{BC}_T(U), X \subseteq Y\}.
\end{aligned}
$$

This definition may be considered as a generalization of equivalence class based definition. We used the successor neighborhood xT to substitute the equivalence class $[x]_E$ in defining one operator and derive the other by duality. If $[x]_E$ is replaced by xT in defining both operators, we would have obtained a pair of operators which are not dual to each other [47]. For a binary relation R, if definition (**def2a**) is used, the set of fixed points of $\overline{apr}_R$ is closed under finite union. This is obviously not required by a closure system. Therefore, we may not use (**def2a**) to establish connections between a closure operator and a binary relation. By generalizing (**def3a**), we have:

$$
\begin{aligned}
\textbf{(def3b)} \qquad \underline{apr}_R(X) &= \bigcup\{xR \mid x \in U, xR \subseteq X\} \\
&= \bigcup\{Y \mid Y \in \mathcal{BO}_R(U), Y \subseteq X\}, \\
\overline{apr}_R(X) &= \bigcap\{\neg xR \mid x \in U, X \subseteq \neg xR\}, \\
&= \bigcap\{Y \mid Y \in \mathcal{BC}_R(U), X \subseteq Y\},
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{BO}_R(U) &= \{xR \mid x \in U\}, \\
\mathcal{BC}_R(U) &= \{\neg xR \mid x \in U\}. \qquad (4.18)
\end{aligned}
$$

Operator $\underline{apr}_R$ satisfies axioms (j1), (j2), and (j3), and operator $\overline{apr}_R$ satisfies axioms (j1′), (j2′), and (j3′). That is, $\overline{apr}_R$ is a closure operator. The families of their fixed points are:

$$\begin{aligned} \mathcal{O}_R(U) &= \{X \mid \underline{apr}_R(X) = X\}, \\ \mathcal{C}_R(U) &= \{X \mid \overline{apr}_R(X) = X\}. \end{aligned} \quad (4.19)$$

The system $\mathcal{C}_R(U)$ is a closure system satisfying axioms (j1), (j2), and (j3). For $x \in U$, the complement of its the successor neighborhood $\neg xR$ is a member of $\mathcal{C}_R(U)$. Any member of $\mathcal{C}_R(U)$ can be expressed as an intersection of a subsystem of $\mathcal{BC}_R(U)$. That is, by finding the intersection closure of the system $\mathcal{BC}_R(U)$ we can derive a closure system $\mathcal{C}_R(U)$. Similar connection can be established between $\mathcal{BO}_R(U)$ and $\mathcal{O}_R(U)$.

Chapter 5

CONCLUSION

This thesis reviews and summaries many existing results on the constructive methods of rough set theory. It contributes to the field by providing a unified framework, a more complete, comprehensive, and coherent study of the problem, and a systematic investigation of the connection established between algebraic methods and constructive methods, and between different constructive methods. The results bring new insights into the theory of rough sets.

5.1 Summary

In Chapter 1, we use the classical rough set model to lead out brief discussion of constructive studies of rough set theory. Three definitions of approximation operators are examined. They are element based, equivalence based, and subsystem based. Properties of rough set approximations are described. Based on such properties, algebraic approaches of rough set theory is discussed and the connections between constructive and algebraic methods are established. The main contents of the thesis include two parts. One part, consisting of Chapters 2 and 3, focuses on two constructive methods for developing generalized rough set models. One is based on binary relation, and the other is based on subsystems. In both approaches, a pair of lower and upper approximation operators is defined. The second part, consisting of Chapter 4, connections between algebraic and constructive approaches, and between different constructive approaches are studied in detail. There is a nice correspondence between properties of binary relations and rough set approximation operators. The classical rough set model defined by an equivalence relation is same as the one defined by a Pawlak topology. The rough set model based on a reflexive and transitive relation equals to the one defined by a topological space. On the other hand, there is a different type of correspondence between rough set models and closure systems.

5.2 Main Contributions

An advantage of constructive approaches is that every notion has a clear and well understood physical meaning. The constructive approach is more suitable for practical applications of rough sets. Many studies have focused on the constructive approach, due to its simplicity and associated intuitiveness. This thesis make additional contributions to the constructive approaches to rough set theory. It presents a more complete, coherent and systematic study. Within the presented framework of Chapter 1, results from existing studies are clearly classified and put into proper perspective. It is shown that the consideration of different but equivalent definitions leads to very different generalization. The connections established between algebraic and constructive methods provide more insights into the theory. Moreover, the new results of such connections fill in a gap in the existing studies.

5.3 Future Research

The results of the thesis have several implications and immediately offer new research problems. One needs to consider other definitions approximation operators, which may introduce more generalized rough set models. One need also to study in more details about the connections between different approaches of rough sets. In this thesis, we only focus on the rough set theory on its own. It is important to study various approaches and notions in relation to other theories, such as modal logic, machine learning, and data analysis. To truly appreciate the usefulness of a theory, one need to find many applications of the theory. It may be useful to investigate the methods and notions discussed in this thesis in real world applications.

Bibliography

[1] Birkhoff, G., *Lattice Theory*, American Mathematical Society, Providence, 1967.

[2] Cattaneo, G., Abstract approximation spaces for rough theories, in: *Rough Sets in Data Mining and Knowledge Discovery*, Polkowski, L. and Skowron, A. (Eds.), Physica Verlag, Berlin, pp. 59-98, 1998.

[3] Cattaneo, G., Mathematical foundations of roughness and fuzziness, manuscript, Dipartimento di Science dell'Informazione, Univerita di Milano, via Comelico 39, I-20135 Milano, Italy, 1997. A preliminary version appeared in *Proceeding of the Fourth International Workshop on Rough Sets, Fuzzy Sets, and Manchine Discovery, RSFD'96*, pp. 241-247, 1996.

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

[5] Chuchro, M., On rough sets in topological Boolean algebras, in: *Rough Sets, Fuzzy Sets and Knowledge Discovery* Ziarko, W.P (Ed.), Springer-Verlag, London, pp. 157-166, 1994.

[6] Cohn, P.M., *Universal Algebra*, Harper and Row Publishers, New York, 1965.

[7] Comer, S., An algebraic approach to the approximation of information, *Fundamenta Informaticae*, 14: 492-502, 1991.

[8] Comer, S., On connections between information systems, rough sets, and algebraic logic, in: *Algebraic Methods in Logic and Computer Science*, Banach Center Publications 28, pp. 117-124, 1993.

[9] Düntsch, I., A logic for rough sets, *Theoretical Computer Science*, 179: 427-, 1997.

[10] Gehrke, M. and Walker, E., On the structure of rough sets, *Bulletin of Polish Academy of Sciences, Mathematics*, 40: 235-245, 1992.

[11] Haack, S., *Philosophy of Logics*, Cambridge University Press, Cambridge, 1978.

[12] Henkin, L., Monk, J.D. and Tarski, A., *Cylindric Algebras I*, North-Holland, Amsterdam, 1971.

[13] Hughes, G.E. and Cresswell, M.J., *An Introduction to Modal logic*, Methuen, London, 1968.

[14] Inuiguchi, M., Hirano, S. and Tsumoto, S. (Eds.), *Rough Set Theory and Granular Computing*, Springer, Berlin, 2003.

[15] Iwinski, T.B., Rough orders and rough concepts, *Bulletin of Polish Academy of Sciences, Mathematics*, 36: 187-192, 1988.

[16] Kortelainen, J., On relationship between modified sets, topological spaces and rough sets, *Fuzzy Sets and Systems*, 61: 91-95, 1994.

[17] 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.

[18] Lin, T.Y., Topological and fuzzy rough sets, in: *Intelligent decision support: Handbook of applications and advances of rough set theory*, Slowinski, R. (Ed.), Kluwer, Dordrecht, pp. 287?04, 1992.

[19] Lin, T.Y. and Liu, Q., Rough approximation operators: axiomatic rough set theory, in: *Rough Sets, Fuzzy Sets and Knowledge Discovery*, Ziarko, W.P. (Ed.), Springer-Verlag, London, pp. 256-260, 1994.

[20] Lin, T.Y., Yao, Y.Y. and Zadeh, L.A. (Eds.), *Data Mining, Rough Sets and Granular Computing*, Physica-Verlag, Heidelberg, 2002.

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

[22] McKinsey, C.C. and Tarski, A., On closed elements in closure algebras, *Annals of Mathematics*, 47: 122-162, 1946.

[23] Pal, S.K., Polkowski, L., and Skowron, A. (Eds.), *Rough-Neural Computing: Techniques for Computing with Words*, Springer, Berlin, 2003.

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

[25] Pawlak, Z., Rough sets, *International Journal of Computer and Information Science*, 11: 341-356, 1982.

[26] Pawlak, Z., Rough sets and fuzzy sets, *Fuzzy Sets system*, 17: 99-102, 1985.

[27] Pawlak, Z. *Rough Sets, Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, 1991.

[28] Polkowski, L. and Skowron, A. (Eds.), *Rough Sets in Knowledge Discovery I, II*, Physica-Verlag, Heidelberg, 1998.

[29] Pomykala, J.A., Approximation operations in approximation space, *Bulletin of the Polish Academy of Sciences, Mathematics*, 35: 653-662, 1987.

[30] Pomykala, J.A., On definability in the nondeterministic information system. *Bulletin of the Polish Academy of Sciences, Mathematics*, 36: 193-210, 1988.

[31] Prior, A.N., Tense logic and the continuum of time, *Studia Logica*, 13: 133-148, 1962.

[32] Rasiowa, H., *An Algebraic Approach to Non-classical Logics*, North-Holland, Amsterdam, 1974.

[33] Slowinski, R., Strict and weak indiscernibility of objects described by quantitative attributes with overlapping norms, *Foundations of Computer and Decision Sciences*, 18: 361-369, 1993.

[34] Slowinski, R. and Vanderpooten, D., Similarity relation as a basis for rough approximations, in: *Advances in Machine Intelligence & Soft-Computing*, Wang, P.P. (Ed.), Department of Electrical Engineering, Duke University, Durham, North Carolina, 17-33, 1997.

[35] Wasilewska, A., Topological rough algebras, in: *Rough Sets and Data Mining: Analysis for Imprecise Data*, Lin, T.Y. and Cercone, N. (Eds.), Kluwer Academic Publishers, Boston, pp. 411-425, 1997.

[36] Wiweger, A., On topological rough sets, *Bulletin of Polish Academy of Sciences, Mathematics*, 37: 89-93, 1989.

[37] Wong, S.K.M. and Yao, Y.Y., Roughness theory, in: *Proceedings of the Sixth International Conference on Information Processing and Management of Uncertainly in Knowledge-Based Systems*, Granada, Spain, pp. 877-883, July 1-5, 1996.

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

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

[40] Yao, Y.Y., A comparative study of fuzzy sets and rough sets. *Information Sciences*, Vol. 109, No. 1-4, pp. 227-242, 1998.

[41] Yao, Y.Y., Combination of rough and fuzzy sets based on α-level sets, in: *Rough Sets and Data Mining: Analysis for Imprecise Data*, Lin, T.Y. and Cercone, N. (Eds.), Kluwer Academic Publishers, Boston, pp. 301-321, 1997.

[42] Yao, Y.Y., Constructive and Algebraic Methods of the Theory of Rough Sets, *Information Sciences*, Vol. 109, No. 1-4, pp. 21-47, 1998.

[43] Yao, Y.Y., Generalized rough set models, in: *Rough Sets in Knowledge Discovery*, Polkowski, L. and Skowron, A. (Eds.), Physica-Verlag, Heidelberg, pp. 286-318, 1998.

[44] Yao, Y.Y., Information granulation and rough set approximation, *International Journal of Intelligent Systems*, Vol. 16, No. 1, pp. 87-104, 2001.

[45] Yao, Y.Y., On generalizing Pawlak approximation operators, in: *Rough Sets and Current Trends in Computing, Proceedings of the First International Conference, RSCTC'98, Lecture Notes in Artificial Intelligence 1424*, Warsaw, Poland, June

22-26, Polkowski, L. and Skowron, A. (Eds.), Springer-Verlag, Berlin, pp. 298-307, 1998.

[46] Yao, Y.Y., On generalizing rough set theory, in: *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Proceedings of the 9th International Conference, RSFDGrC 2003, Lecture Notes in Artificial Intelligence 2639*, Chongqing, China, May 26-29, Wang, G., Liu, Q., Yao, Y. and Skowron, A. (Eds.), pp. 44-51, 2003.

[47] Yao, Y.Y., Relational interpretations of neighborhood operators and rough set approximation operators, *Information Sciences*, Vol. 111, No. 1-4, pp. 239-259, 1998.

[48] Yao, Y.Y., Two views of the theory of rough sets in finite universes. *International Journal of Approximate Reasoning*, 15: 291-317, 1996.

[49] Yao, Y.Y. and Li, X., Comparison of rough-set and interval-set models for uncertain reasoning, *Fundamenta Informaticae*, Vol. 27, No. 2-3, pp. 289-298, 1996.

[50] Yao, Y.Y. and Lin, T.Y., Generalization of Rough Sets using Modal Logic, *Intelligent Automation and Soft Computing, an International Journal*, 2: 103-120, 1996.

[51] Yao, Y.Y. and Lingras, P.J., Interpretations of belief functions in the theory of rough sets, *Information Sciences*, Vol.104, No. 1-2, pp. 81-106, 1998.

[52] Yao, Y.Y. and Zhong, N., Granular computing using information tables, in: *Data Mining, Rough Sets and Granular Computing*, Lin, T.Y., Yao, Y.Y. and Zadeh, L.A. (Eds.), Physica-Verlag, Heidelberg, pp. 102-124, 2002.

[53] Yao, Y.Y., Wong, S.K.M., and Lin, T.Y., A review of rough set models, in: *Rough Sets and Data Mining: Analysis for Imprecise Data*, Lin, T.Y. and Cercone, N. (Eds.), Kluwer Academic Publishers, Boston, pp. 47-75, 1997.

[54] Zadeh, L.A., *Fuzzy sets.* Information Contribution. 8: 338-353, 1965.

[55] Zakowski, W., Approximations in the space (U,II), *Demonstratio Mathematica*, XVI: 761-769, 1983.