Rough Set Approximations in Formal Concept Analysis

Y.Y. Yao, Yaohua Chen

Department of Computer Science, University of Regina
Regina, Saskatchewan, Canada S4S 0A2
E-mail: {yyao, chen115y}@cs.uregina.ca

***Abstract*— An important topic of rough set theory is the approximation of undefinable sets or concepts through definable sets. It involves the construction of a system of definable sets and the definition of approximation operators. In this paper, the notion of rough set approximations is introduced into formal concept analysis. Approximation operators are defined based on both lattice-theoretic and set-theoretic operators. The results provide a better understanding of data analysis using rough set theory and formal concept analysis.**

I. Introduction

An underlying notion of rough set theory is the indiscernibility of objects [8], [9]. By modelling indiscernibility as an equivalence relation, one can partition a finite universe of objects into pair-wise disjoint subsets. The partition provides a granulated view of the universe. An equivalence class is considered as a whole, instead of many individuals. In other words, one can only observe, measure, or characterize equivalence classes. The empty set, equivalence classes and unions of equivalence classes form a system of definable subsets under indiscernibility. All subsets not in the system are consequently approximated through definable sets.

Formal concept analysis is developed based on a formal context, which is a binary relation between a set of objects and a set of attributes or properties. From a formal context, one can construct (objects, properties) pairs known as the formal concepts [4], [11]. The set of objects is referred to as the extension, and the set of properties as the intension, of a formal concept. They uniquely determine each other. The family of all formal concepts is a complete lattice. The extension of a formal concept can be viewed a definable set of objects, although in a sense different from that of rough set theory [16], [17].

A comparative examination of rough set theory and formal concept analysis shows that each of them deals with a particular type of definability. The common notion of definability links the two theories together. One can immediately adopt ideas from each other [16], [17]. The notions of formal concept and formal concept lattice can be introduced into rough set theory by considering different types of formal concepts [17]. Rough set approximation operators can be introduced into formal concept analysis by considering a different type of definability [5]. The combination of the two theory would produce new tools for data analysis.

Two types of formulations have been suggested by some authors for introducing concept approximations into formal concept analysis. One is the introduction of an equivalence relation on the set of objects and/or the set of properties [6], the other is the use of the system of definable concepts in the concept lattice [5], [10].

Kent argued that rough set theory and formal concept analysis have much in common, in terms of goals and methodologies [6]. A framework of rough formal concept analysis was introduced as a synthesis of the two theories [6]. Kent's formulation is based on an equivalence relation on the set of objects. With respect to the formal context, a pair of lower and upper contextual approximations is first defined. The two contextual approximations are then used to define a pair of lower and upper approximations of formal concepts.

Saquer and Deogun studied approximations of a set of objects, a set of properties, and a pair of a set of objects and a set of properties, based on the system of formal concepts in the concept lattice [10]. For example, given a set of objects, they attempted to approximate it by formal concepts whose extensions approximate the set. An equivalence relation is introduced on the set of objects from a formal context, which leads to rough set approximations. However, their formulation is slightly flawed and fails to achieve such a goal. An equivalence class is not necessarily the extension of a formal concept. The union of extensions of a family of formal concepts may not be the extension of a formal concept. Consequently, as pointed out by Hu *et al.* [5], the approximations defined by Saquer and Deogun may not necessarily be formal concepts.

Hu *et al.* suggested an alternative formulation to ensure that approximations are indeed formal concepts [5]. Instead of defining an equivalence relation, they defined a partial order on the set of objects. For an object, its principal filter, which is the set of objects "greater than or equal to" the object and is called the partial class by Hu *et al.*, is the extension of a formal concept. The family of all principal filters is the set of join-irreducible elements of the concept lattice. Similarly, a partial order relation can be defined on the set of properties. The family of meet-irreducible elements of the concept lattice can be constructed. However, their definition of lower approximation based on the union of extensions of formal concept has the same shortcoming of Saquer and Deogun's definition [10].

Based on the common notion of definability and the results from the above studies, we propose a framework to examine the issues of rough set approximations within formal concept analysis. We concentrate on the interpretations and formulations of various notions, instead of efficient algorithms for constructing approximations. It is shown that the problem with existing studies can be easily solved by a clear separation of two systems, the formal concept lattice and the system of extensions of formal concepts. The two systems give rise to two different types of approximations.

II. Subsystem Based Formulation of Rough Set Theory

The rough set theory is an extension of classical set theory with two additional approximation operators [12]. Various formulations of the theory have been proposed and studied [13], [14], [15]. In the subsystem based formulation, a subsystem of the power set of a universe is first constructed and the approximation operators are then defined using the subsystem.

Suppose U is a finite and nonempty universe of objects. Let $E \subseteq U \times U$ be an equivalence relation on U. The equivalence relation divides the universe into a family of pair-wise disjoint subsets, called the partition of the universe and denoted by U/E. The pair $apr = (U, E)$ is referred to as an approximation space.

An approximation space induces a granulated view of the universe. For an object $x \in U$, the equivalence class containing x is given by:

$$[x]_E = \{y \mid xEy\}. \tag{1}$$

Intuitively speaking, objects in $[x]_E$ are indistinguishable from x. One is therefore forced to consider $[x]_E$ as a whole. In other words, under an equivalence relation, $[x]_E$s are the smallest non-empty observable, measurable, or definable subsets of 2^U, the power set of U. By extending the definability of equivalence classes, we assume that a union of some equivalence classes is also definable. The family of definable subsets contains the empty set $\emptyset$ and is closed under set complement, intersection, and union. It is an σ-algebra whose basis is U/E. Let $\sigma(U/E) \subseteq 2^U$ denote the subsystem of definable sets of objects.

A subset of objects not in $\sigma(U/E)$ is said to be undefinable. An undefinable set must be approximated from below and above by a pair of definable sets.

Definition 1: In an approximation space $apr = (U, E)$, a pair of approximation operators, $\underline{apr}$, $\overline{apr} : 2^U \longrightarrow 2^U$, are defined by:

$$\begin{aligned} \underline{apr}(A) &= \bigcup\{X \mid X \in \sigma(U/E), X \subseteq A\}, \\ \overline{apr}(A) &= \bigcap\{X \mid X \in \sigma(U/E), A \subseteq X\}. \end{aligned} \tag{2}$$

The lower approximation $\underline{apr}(A) \in \sigma(U/E)$ is the greatest definable set contained in A, and the upper approximation $\overline{apr}(A) \in \sigma(U/E)$ is the least definable set containing A.

The lower and upper approximation operators have the following properties: for two sets of objects A and B,

$$\begin{aligned}
&\text{(i).} && \underline{apr}(A) = (\overline{apr}(A^c))^c, \\
& && \overline{apr}(A) = (\underline{apr}(A^c))^c; \\
&\text{(ii).} && \underline{apr}(A \cap B) = \underline{apr}(A) \cap \underline{apr}(B), \\
& && \overline{apr}(A \cup B) = \overline{apr}(A) \cup \overline{apr}(B); \\
&\text{(iii).} && \underline{apr}(A) \subseteq A \subseteq \overline{apr}(A); \\
&\text{(iv).} && \underline{apr}(\underline{apr}(A)) = \underline{apr}(A), \\
& && \overline{apr}(\overline{apr}(A)) = \overline{apr}(A); \\
&\text{(v).} && \underline{apr}(\overline{apr}(A)) = \overline{apr}(A), \\
& && \overline{apr}(\underline{apr}(A)) = \underline{apr}(A).
\end{aligned}$$

Property (i) states that the approximation operators are dual operators with respect to set complement ^{c}. Property (ii) states that the lower approximation operator is distributive over set intersection $\cap$, and the upper approximation operator is distributive over set union $\cup$. By property (iii), a set lies within its lower and upper approximations. Properties (iv) and (v) deal with the compositions of lower and upper approximation operators. The result of the composition of a sequence of lower and upper approximation operators is the same as the application of the approximation operator closest to A.

The approximation operators can also be defined by using equivalence classes.

Definition 2: In an approximation space $apr = (U, E)$, a pair of approximation operators, $\underline{apr}$, $\overline{apr} : 2^U \longrightarrow 2^U$, are defined by:

$$\begin{aligned} \underline{apr}(A) &= \bigcup\{[x]_E \mid [x]_E \in U/E, [x]_E \subseteq A\}, \\ \overline{apr}(A) &= \bigcup\{[x]_E \mid [x]_E \in U/E, A \cap [x]_E \neq \emptyset\}. \end{aligned} \tag{3}$$

That is, the lower approximation is the union of equivalence classes that are subsets of A, and the upper approximation is the union of equivalence classes that have a non-empty intersection with A.

As shown by the following theorem, the approximation operators truthfully reflect our intuitive understanding of the notion of definability.

Theorem 1: In an approximation space $apr = (U, E)$, for a set of objects A, $\underline{apr}(A) = \overline{apr}(A)$ if and only if $A \in \sigma(U/E)$.

An important implication of the theorem is that for an undefinable set $A \subseteq U$ we have $\underline{apr}(A) \neq \overline{apr}(A)$. In fact, $\underline{apr}(A)$ is a proper subset of $\overline{apr}(A)$, namely, $\underline{apr}(A) \subset \overline{apr}(A)$.

The basic ideas of subsystem based formulation can be generalized by considering different subsystems that represent different types of definability [13]. In the next section, we demonstrate that formal concept analysis can be used to construct subsystem of definable sets. One can immediately apply the subsystem based formulation into formal concept analysis and introduce approximation operators.

III. Formal Concept Analysis

Formal concept analysis deals with visual presentation and analysis of data [4], [11]. It focuses on the definability of a set of objects based on a set of properties, and vice versa.

Let U and V be any two finite sets. Elements of U are called objects, and elements of V are called properties. The relationships between objects and properties are described by a binary relation R between U and V, which is a subset of the Cartesian product $U \times V$. For a pair of elements $x \in U$ and $y \in V$, if $(x, y) \in R$, written as xRy, x has the property y, or the property y is possessed by object x. The triplet (U, V, R) is called a formal context.

Based on the binary relation, we associate a set of properties to an object. An object $x \in U$ has the set of properties:

$$xR = \{y \in V \mid xRy\} \subseteq V. \tag{4}$$

Similarly, a property y is possessed by the set of objects:

$$Ry = \{x \in U \mid xRy\} \subseteq U. \tag{5}$$

By extending these notations, we can establish relationships between subsets of objects and subsets of properties. This leads to two operators, one from 2^U to 2^V and the other from 2^V to 2^U.

Definition 3: Suppose (U, V, R) is a formal context. For a subset of objects, we associate it with a set of properties:

$$\begin{aligned} X^* &= \{y \in V \mid \forall x \in U(x \in X \Longrightarrow xRy)\} \\ &= \{y \in V \mid X \subseteq Ry\} \\ &= \bigcap_{x \in X} xR. \end{aligned} \tag{6}$$

For a subset of properties, we associate it with a set of objects:

$$\begin{aligned} Y^* &= \{x \in U \mid \forall y \in V(y \in Y \Longrightarrow xRy)\} \\ &= \{x \in U \mid Y \subseteq xR\} \\ &= \bigcap_{y \in Y} Ry. \end{aligned} \tag{7}$$

For simplicity, the same symbol is used for both operators. The actual role of the operators can be easily seen from the context.

By definition, $\{x\}^* = xR$ is the set of property possessed by x, and $\{y\}^* = Ry$ is the set of objects having property y. For a set of objects X, X^* is the *maximal* set of properties shared by *all* objects in X. For a set of properties Y, Y^* is the *maximal* set of objects that have *all* properties in Y.

The operators * have the following properties: for $X, X_1, X_2 \subseteq U$ and $Y, Y_1, Y_2 \subseteq V$,

(1) $X_1 \subseteq X_2 \Longrightarrow X_1^* \supseteq X_2^*$,
$Y_1 \subseteq Y_2 \Longrightarrow Y_1^* \supseteq Y_2^*$,

(2) $X \subseteq X^{**}$,
$Y \subseteq Y^{**}$,

(3) $X^{***} = X^*$,
$Y^{***} = Y^*$,

(4) $(X_1 \cup X_2)^* = X_1^* \cap X_2^*$,
$(Y_1 \cup Y_2)^* = Y_1^* \cap Y_2^*$.

In formal concept analysis, one is interested in a pair of a set of objects and a set of properties that uniquely define each other. More specifically, for $(X, Y) = (Y^*, X^*)$, we have:

$$\begin{aligned} x \in X &\Longleftrightarrow x \in Y^* \\ &\Longleftrightarrow Y \subseteq xR \\ &\Longleftrightarrow \bigwedge_{y \in Y} xRy; \end{aligned} \tag{8}$$

$$\begin{aligned} \bigwedge_{x \in X} xRy &\Longleftrightarrow X \subseteq Ry \\ &\Longleftrightarrow y \in X^* \\ &\Longleftrightarrow y \in Y. \end{aligned} \tag{9}$$

That is, the set of objects X is defined based on the set of properties Y, and vice versa. This type of definability leads to the introduction of the notion of formal concepts.

Definition 4: A pair (X, Y), $X \subseteq U$, $Y \subseteq V$, is called a formal concept of the context (U, V, R), if $X = Y^*$ and $Y = X^*$. Furthermore, $X = ex(X, Y)$ is called the extension of the concept, and $Y = in(X, Y)$ is the intension of the concept.

The set of all formal concepts forms a complete lattice called a concept lattice, denoted by $L(U, V, R)$ or simply L. The meet and join of the lattice is characterized by the following basic theorem of concept lattices [4], [11].

Theorem 2: The formal concept lattice L is a complete lattice in which the meet and join are given by:

$$\begin{aligned} \bigwedge_{t \in T} (X_t, Y_t) &= (\bigcap_{t \in T} X_t, (\bigcup_{t \in T} Y_t)^{**}), \\ \bigvee_{t \in T} (X_t, Y_t) &= ((\bigcup_{t \in T} X_t)^{**}, \bigcap_{t \in T} Y_t). \end{aligned} \tag{10}$$

where T is an index set and for every $t \in T$, (X_t, Y_t) is a formal concept.

The order relation of the lattice can be defined based on the set inclusion relation.

Definition 5: For two formal concepts (X_1, Y_1) and (X_2, Y_2), (X_1, Y_1) is a sub-concept of (X_2, Y_2), written $(X_1, Y_1) \preceq (X_2, Y_2)$, and (X_2, Y_2) is a super-concept of (X_1, Y_1), if and only if $X_1 \subseteq X_2$, or equivalently, if and only if $Y_2 \subseteq Y_1$.

A more general (specific) concept is characterized by a larger (smaller) subset of objects that share a smaller (larger) subset of properties.

The lattice-theoretic operators of meet ($\wedge$) and join ($\vee$) of the concept lattice are defined based on the set-theoretic operators of intersection ($\cap$) and union ($\cup$). However, they are not the same. An intersection of extensions (intensions) of a family of formal concepts is the extension (intension) of a formal concept. A union of extensions (intensions) of a family of formal concepts is not necessarily the extension (intension) of a formal concept. Given two formal concepts,

TABLE I

A FORMAL CONTEXT TAKEN FROM [4]

	a	b	c	d	e	f	g	h	i
1. Leech	×	×					×		
2. Bream	×	×					×	×	
3. Frog	×	×	×				×	×	
4. Dog	×		×				×	×	×
5. Spike-weed	×	×		×		×			
6. Reed	×	×	×	×		×			
7. Bean	×		×	×	×				
8. Maize	×		×	×		×			

a: needs water to live; b: lives in water; c: lives on land; d: needs chlorophyll to produce food; e: two seed leaves; f: one seed leaf; g: can move around; h: has limbs; i: suckles its offspring.

one can find the extension of their meet by the set intersection of their extensions, and the intension of their join by the set intersection of their intensions. One cannot find directly the intension of their meet and the extension of their join by simply applying set-theoretic operators.

Example 1: The ideas of formal concept analysis can be illustrated by an example. Table I gives a formal context and Figure 1 gives the corresponding concept lattice. Consider two formal concepts $(\{3,6\},\{a,b,c\})$ and $(\{5,6,7,8\},\{a,d\})$. Their meet is the formal concept:

$$(\{3,6\}\cap\{5,6,7,8\},(\{a,b,c\}\cup\{a,d\})^{**}) = (\{6\},\{a,b,c,d,f\}),$$

and their join is the formal concept:

$$((\{3,6\}\cup\{5,6,7,8\})^{**},\{a,b,c\}\cap\{a,d\}) = (\{1,2,3,4,5,6,7,8\},\{a\}).$$

The intersection of extensions of two concepts is the extension of their meet, and the intersection of the intensions is the intension of their join. On the other hand, the union of extensions of the two concepts is $\{3,5,6,7,8\}$, which is not the extension of any formal concept. The union of the intensions is $\{a,b,c,d\}$, which is not the intension of any formal concept.

IV. APPROXIMATIONS IN FORMAL CONCEPT ANALYSIS

A formal concept consists of a definable set of objects and a definable set of properties. The concept lattice is the family of all such definable concepts. Given an arbitrary set of objects, it may not be the extension of a formal concept. The set can therefore be viewed as an undefinable set of objects. Following the theory of rough sets, such a set of objects can be approximated by definable sets of objects, namely, the extensions of formal concepts. In this section, two methods of approximations are discussed by using the subsystem based formulation of rough set theory.

A. Approximations based on lattice-theoretic operators

Hu *et al.* suggested a method for approximation using ideas similar to the subsystem based formulation of rough set theory [5]. The concept lattice is used as the system of

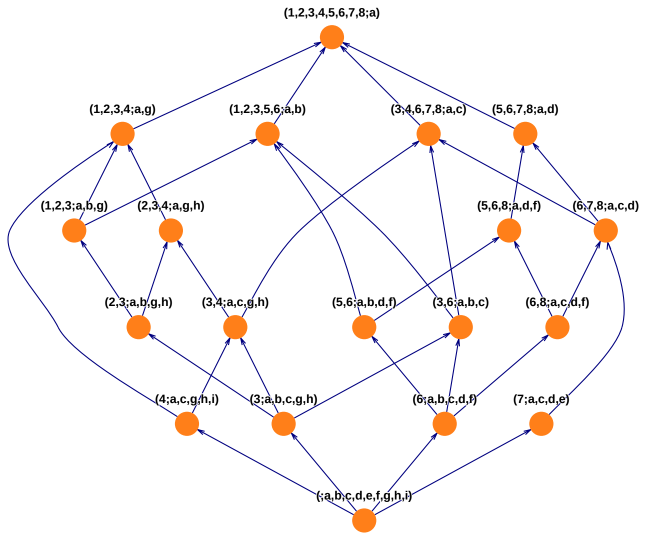

Fig. 1. Concept lattice for the context of Table 1, produced by "Formal Concept Calculator" (developed by Sören Auer, http://www.advis.de/soeren/fca/).

definable concepts, and lattice-theoretic operators are used to define approximation operators. We present a modified formulation of their method by fixing its minor flaw.

For a subset of objects $A \subseteq U$, suppose we want to approximate it by the extensions of a pair of formal concepts in the concept lattice. We can extend Definition 1 to achieve this goal. In equation (2), set-theoretic operators $\cap$ and $\cup$ are replaced by lattice-theoretic operators $\wedge$ and $\vee$, the subsystem $\sigma(U/E)$ by lattice L, and definable set of objects by formal concepts. The extensions of the resulting two concepts are the approximations of A.

Definition 6: For a subset of objects $A \subseteq U$, its lower and upper approximations are defined by:

$$\begin{aligned} \underline{lapr}(A) &= ex(\bigvee\{(X,Y)\in L \mid X\subseteq A\}), \\ \overline{lapr}(A) &= ex(\bigwedge\{(X,Y)\in L \mid A\subseteq X\}). \end{aligned} \quad (11)$$

The lower approximation of a set of objects A is the extension of the formal concept $(\underline{lapr}(A),(\underline{lapr}(A))^*)$, and the upper approximation is the extension of the formal concept $(\overline{lapr}(A),(\overline{lapr}(A))^*)$. The concept $(\underline{lapr}(A),(\underline{lapr}(A))^*)$ is the supremum of those concepts whose extensions are subsets of A, and the concept $(\overline{lapr}(A),(\overline{lapr}(A))^*)$ is the infimum of those concepts whose extensions are supersets of A.

For a formal concept (X,Y), X^c may not necessarily be the extension of a formal concept. The concept lattice in general is not a complemented lattice. The approximation operators $\underline{lapr}$ and $\overline{lapr}$ are not dual operators.

Recall that an intersection of extensions is an extension of a concept, but the union of extensions may not be the extension of a concept. It follows that $(\overline{lapr}(A),(\overline{lapr}(A))^*)$ is the smallest concept whose extension is a superset of A. However, $(\underline{lapr}(A),(\underline{lapr}(A))^*)$ may not be the largest concept whose extension is a subset of A. The new approximation operators do not satisfy property (ii). They only satisfy a week version known as monotonicity with respect to set inclusion:

$$\begin{array}{ll} \text{(vi)} & A \subseteq A' \Longrightarrow \underline{lapr}(A) \subseteq \underline{lapr}(A'), \\ & A \subseteq A' \Longrightarrow \overline{lapr}(A) \subseteq \overline{lapr}(A'). \end{array}$$

By property (2), for a family of concepts (X_t, Y_t), we have $\bigcup_{t \in T} X_t \subseteq (\bigcup_{t \in T} X_t)^{**}$. Thus, although $X_t \subseteq A$ for all $t \in T$, it may happen that $A \subseteq \underline{lapr}(A)$. That is, the lower approximation of A may not be a subset of A. With respect to property (iii), we have a weaker version:

$$\begin{array}{ll} \text{(vii)} & \underline{lapr}(A) \subseteq \overline{lapr}(A), \\ \text{(viii)} & A \subseteq \overline{lapr}(A). \end{array}$$

Both $\underline{lapr}(A)$ and $\overline{lapr}(A)$ are extensions of formal concepts. It follows that the operators $\underline{lapr}$ and $\overline{lapr}$ satisfy properties (iv) and (v).

Example 2: Given the concept lattice in Figure 1, consider a set of objects $A = \{3, 5, 6\}$. The family of subsets of A that are extensions of concepts is:

$$\{\ \emptyset, \{3\}, \{6\}, \{3, 6\}, \{5, 6\}\ \}.$$

The corresponding family of concepts is:

$$\begin{array}{l} \{\ (\emptyset, \{a, b, c, d, e, f, g, h, i\}), \\ \quad (\{3\}, \{a, b, c, g, h\}), (\{6\}, \{a, b, c, d, f\}), \\ \quad (\{3, 6\}, \{a, b, c\}), (\{5, 6\}, \{a, b, d, f\})\ \}. \end{array}$$

Their supremum is $(\{1, 2, 3, 5, 6\}, \{a, b\})$. The lower approximation is $\underline{lapr}(A) = \{1, 2, 3, 5, 6\}$, which is indeed a superset of A. The family of supersets of A that are extensions of concepts is:

$$\{\{1, 2, 3, 5, 6\}, \{1, 2, 3, 4, 5, 6, 7, 8\}\}.$$

The corresponding family of concepts is:

$$\{\ (\{1, 2, 3, 5, 6\}, \{a, b\}), (\{1, 2, 3, 4, 5, 6, 7, 8\}, \{a\})\ \}.$$

Their infimum is $(\{1, 2, 3, 5, 6\}, \{a, b\})$. The upper approximation is $\overline{lapr}(A) = \{1, 2, 3, 5, 6\}$, which is the smallest concept whose extension contains A. Although A is not an extension of a concept, it has the same lower and upper approximations.

In contrast to Theorem 1, we can only obtain a weak version.

Theorem 3: In a concept lattice $L(U, V, R)$, if A is an extension of a concept, i.e., (A, A^*) is a concept, then $\underline{lapr}(A) = \overline{lapr}(A)$.

As shown by the example, the reverse implication in the theorem is not true. This is a shortcoming of the formulation based on lattice-theoretic operators.

Hu *et al.* proposed another equivalent definition of approximation operators by considering, in Definition 6, only the families of meet irreducible and join irreducible formal concepts [5]. Their definition is similar to the idea of Definition 2.

The upper approximation operator $\overline{lapr}$ is related to the operator *. For any set of objects $A \subseteq U$, we can derive a set of properties A^*. For the set of properties A^*, we can derive another set of objects A^{**}. By property (3), (A^{**}, A^*) is a formal concept. By property (2), we have $A \subseteq A^{**}$. If fact, (A^{**}, A^*) is the smallest formal concept whose extension contains A. This offers another definition of upper approximation operator.

Definition 7: For a subset of objects $A \subseteq U$, its upper approximation is defined by:

$$\overline{lapr}(A) \quad = \quad A^{**}. \tag{12}$$

The idea of approximating a set of objects can be used to define operators that approximate a set of properties.

Definition 8: For a subset of properties $B \subseteq V$, its lower and upper approximations are defined by:

$$\begin{aligned} \underline{lapr}(B) &= in(\bigwedge\{(X, Y) \in L \mid B \subseteq Y\}), \\ \overline{lapr}(B) &= in(\bigvee\{(X, Y) \in L \mid Y \subseteq B\}). \end{aligned} \tag{13}$$

The lower approximation of a set of properties B is the intension of the formal concept $((\underline{lapr}(B))^*, \underline{lapr}(B))$, and the upper approximation is the intension of the formal concept $((\overline{lapr}(B))^*, \overline{lapr}(B))$.

B. Approximations based on set-theoretic operators

By comparing with the standard rough set approximations, one can observe two problems of the approximation operators defined based on lattice-theoretic operators. The lower approximation of a set of objects A is not necessary a subset of A. Although a set of objects A is undefinable, i.e., A is not the extension of a formal concept, its lower and upper approximations may be the same. In order to avoid these shortcomings, we present another formulation by using set-theoretic operators.

The extension of a formal concept is a definable set of objects. A system of definable sets can be derived from a concept lattice.

Definition 9: For a formal concept lattice L, the family of all extensions is given by:

$$EX(L) = \{ex(X, Y) \mid (X, Y) \in L\}. \tag{14}$$

The system $EX(L)$ contains the empty set $\emptyset$, the entire set U, and is closed under intersection. Thus, it is a closure system [1]. Although one can define the upper approximation by extending Definition 1, one can not define the lower approximation. Nevertheless, one can still keep the intuitive interpretations of lower and upper approximations. That is, the lower approximation is a largest set in $EX(L)$ that is contained in A, and the upper approximation is a smallest set in $EX(L)$ that contains A. In this case, while the smallest set containing A is unique, the largest set contained in A is no longer unique.

Definition 10: For a subset of objects $A \subseteq U$, its upper approximation is defined by:

$$\overline{sapr}(A) = \bigcap \{X \mid X \in EX(L), A \subseteq X\}, \tag{15}$$

and its lower approximation is a family of sets:

$$\begin{aligned}\underline{sapr}(A) \;=\; & \{X \mid X \in EX(L), X \subseteq A, \\ & \forall X' \in EX(L)(X \subseteq X' \Longrightarrow X \not\subseteq A)\}.\end{aligned} \tag{16}$$

The upper approximation $\overline{sapr}(A)$ is in fact the same as $\overline{lapr}(A)$, namely, $\overline{sapr}(A) = \overline{lapr}(A)$. However, the lower approximation is different. An important feature is that a set can be approximated from below by several definable sets of objects. In general, for $A' \in \underline{sapr}(A)$, we have $A' \subseteq \underline{lapr}(A)$.

Example 3: In the concept lattice L of Figure 1, the family of all extensions $EX(L)$ are:

$$\begin{aligned} EX(L) \;=\; \{ & \emptyset, \{3\}, \{4\}, \{6\}, \{7\}, \\ & \{2,3\}, \{3,4\}, \{3,6\}, \{5,6\}, \{6,8\}, \\ & \{1,2,3\}, \{2,3,4\}, \{6,7,8\}, \{5,6,8\}, \\ & \{1,2,3,4\}, \{5,6,7,8\}, \\ & \{1,2,3,5,6\}, \{3,4,6,7,8\}, \\ & \{1,2,3,4,5,6,7,8\} \}. \end{aligned}$$

For a set of objects $A = \{3,5,6\}$, the lower approximation is given by $\underline{sapr}(A) = \{\{3,6\},\{5,6\}\}$, which is a family of sets of objects. The upper approximation is given by $\overline{sapr}(A) = \{1,2,3,5,6\}$, which is a unique set of objects.

With respect to property (iii), we have:

(ix) $\quad A' \subseteq A \subseteq \overline{sapr}(A), \quad \text{for all } A \in \underline{sapr}(A).$

That is, A lies within any of its lower approximation and upper approximation. For the set-theoretic formulation, we have a theorem corresponding to Theorem 1.

Theorem 4: In a concept lattice $L(U,V,R)$, for a set of objects A, $\overline{sapr}(A) = A$ and $\underline{sapr}(A) = \{A\}$, if and only if A is an extension of a concept.

In the new formulation, we resolve the difficulties with the approximation operators $\underline{lapr}$ and $\overline{lapr}$. The lower approximation $\underline{sapr}$ offers more insights into the notion of approximations. In some situations, the union of a family of definable sets is not necessarily a definable set. It may not be reasonable to insist on a unique approximation. The approximation of a set by a family of set may provide a better characterization of the set.

V. Conclusion

One of the issues studied in rough set theory is the approximation of undefinable sets through definable sets. Typically, the family of definable sets is a subsystem of the power set of a universe. There are many ways to construct a subsystem of definable sets [13]. Formal concept analysis offers a different approach for the construction of a family of definable sets. The notion of approximations can be introduced naturally into formal concept analysis.

Formal concepts in a formal concept lattice correspond to definable sets. Two types of approximation operators are examined, one is based on lattice-theoretic operators and the other is based on set-theoretic operators. Their properties are investigated. A distinguishing feature of the lower approximation defined by set-theoretic operators is that a set is approximated from below by a family of sets, instead of a unique set as in the standard rough set theory.

The theory of rough sets and formal concept analysis capture different aspects of data. The introduction of the notion of approximations into formal concept analysis combines the two theories, which improves our understanding of data and produces new tools for data analysis.

The derivation operators * is an example of modal-style operators [2], [3], [16]. One can study the notion of rough set approximations in a general framework in which various modal-style operators are defined based on a formal context [2], [3], [7], [16].

References

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

[2] Düntsch, I. and Gediga, G. Approximation operators in qualitative data analysis, in: *Theory and Application of Relational Structures as Knowledge Instruments*, de Swart, H., Orlowska, E., Schmidt, G. and Roubens, M. (Eds.), Springer, Heidelberg, 216-233, 2003.

[3] Gediga, G. and Düntsch, I. Modal-style operators in qualitative data analysis, *Proceedings of the 2002 IEEE International Conference on Data Mining*, 155-162, 2002.

[4] Ganter, B. and Wille, R. *Formal Concept Analysis: Mathematical Foundations*, Springer-Verlag, New York, 1999.

[5] Hu, K., Sui, Y., Lu, Y., Wang, J., and Shi, C. Concept approximation in concept lattice, *Knowledge Discovery and Data Mining, Proceedings of 5th Pacific-Asia Conference, PAKDD'01*, 167-173, 2001.

[6] Kent, R.E. Rough concept analysis, *Rough Sets, Fuzzy Sets and Knowledge Discovery, Proceedings of the International Workshop on Rough Sets and Knowledge Discovery, RSKD'93*, 245–253, 1993.

[7] Pagliani, P. From concept lattices to approximation spaces: algebraic structures of some spaces of partial objects, *Fundamenta Informaticae*, **18**, 1-25, 1993.

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

[9] Pawlak, Z. *Rough Sets - Theoretical Aspects of Reasoning About Data*, Kluwer Publishers, Boston, Dordrecht, 1991.

[10] Saquer, J. and Deogun, J.S. Formal rough concept analysis, *Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing, Proceedings of 7th International Workshop, RSFDGrC '99*, 91-99, 1999.

[11] Wille, R. Restructuring lattice theory: an approach based on hierarchies of concepts, in: *Ordered sets*, Ivan Rival (Ed.), Reidel, Dordecht-Boston, 445-470, 1982.

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

[13] Yao, Y.Y. On generalizing Pawlak approximation operators, *Rough Sets and Current Trends in Computing, Proceedings of 1st International Conference, RSCTC'98*, 298-307, 1998.

[14] Yao, Y.Y. Constructive and algebraic methods of the theory of rough sets, *Information Sciences*, **109**, 21-47, 1998.

[15] Yao, Y.Y. On generalizing rough set theory, *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Proceedings of 9th International Conference, RSFDGrC'03*, 44-51, 2003.

[16] Yao, Y.Y. A comparative study of formal concept analysis and rough set theory in data analysis, *Rough Sets and Current Trends in Computing, Proceedings of 3rd International Conference, RSCTC'04*, 2004.

[17] Yao, Y.Y. Concept lattices in rough set theory, *Proceedings of 23rd International Meeting of the North American Fuzzy Information Processing Society*, 2004.