Concept Analysis Using Quantitative Structured Three-Way Rough Set Approximations

Mengjun Hu[(✉)]

Department of Computer Science, University of Regina,
Regina, SK S4S 0A2, Canada
mengjun.hu@uregina.ca

Abstract. One important topic of concept analysis is to learn an intension of a concept through a given extension. In the case where an exact intension cannot be formulated due to limited information, rough set theory introduces approximations to roughly learn the intension. Pawlak originally proposes a qualitative formulation of approximations which allows no error in the learned intension. Various quantitative formulations have been studied as generalizations, most of which use probabilistic measures. In contrast, non-probabilistic formulations have not been fully investigated. On the other hand, three-way approximations and structured approximations have been proposed to emphasize the semantics of approximations for the purpose of learning and interpreting intension. To combine the benefits of these two directions of generalizations, this paper investigates quantitative structured three-way approximations based on both probabilistic and non-probabilistic measures in the context of both complete and incomplete information.

Keywords: Concept analysis · Three-way · Rough set · Incomplete information · Subsethood measure

1 Introduction

Concept analysis is one common application of rough set theory [18,19]. A concept can be formally represented by a pair of intension and extension [3] where the intension describes the definition and the extension lists all instances. Concept analysis is usually based on a dataset represented in a tabular form with rows as objects and columns as attributes [7,20,21,24,31,32]. The attributes are used to describe the properties of objects as well as to formulate intensions. While learning extension from a given intension is not difficult, the opposite task may be complicated. In particular, we may not be able to find an exact intension of a given extension due to insufficient, incomplete, or limited information.

To solve the above issue, rough set theory introduces the concept of approximations to roughly approximate the true intension. As illustrated by Fig. 1(a) [12], the set of all objects in a given table, represented by the biggest

rectangle, is divided into pieces called definable sets. Each definable set can be precisely described by an intension and is used to approximate a given extension. Pawlak [18,19] proposes a pair of lower and upper approximations, where the lower approximation corresponds to the positive region in Fig. 1(a) and the upper approximation corresponds to the union of the positive and boundary regions. A set of classification rules is derived from an approximation by using the intension of a definable set as the premise of one rule. All such rules together approximate the true intension and can be used to classify instances of the concept. By interpreting classification rules through associated actions, Yao [26] further considers the negative region, which leads to a three-way approximation consisting of the positive, boundary, and negative regions. Semantically, the positive and negative regions are associated with actions of accepting and rejecting instances of the concept, respectively. The boundary region is associated with a non-commitment action, which reflects the limitation of our knowledge.
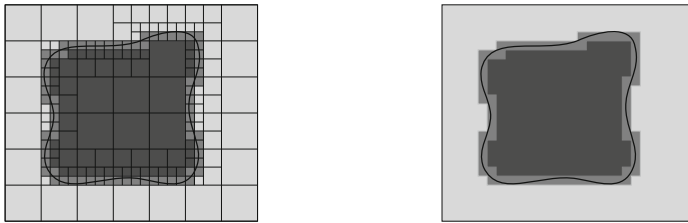
(a) Structured approximations   (b) Unstructured approximations

☐ a definable set   ⬡ a given extension X ■ positive   ■ boundary   ☐ negative

Fig. 1. Illustrations of rough set approximations [12]

There are at least two directions in which the above approximations can be improved. The first direction is to allow a certain rate of misclassification in order to enlarge the positive and negative regions and shrink the boundary region. Instead of the qualitative set-inclusion used in Fig. 1(a), a consideration of quantitative measures results in various quantitative rough set models. Most related research uses probabilistic measures [9,23,29,33] and a few considers non-probabilistic [10,28]. Yao and Deng [28] propose a general framework of formulating both probabilistic and non-probabilistic approximations based on subsethood measures whose properties are further studied in [11]. A few related works regarding concept-based non-probabilistic classifiers are investigated in [17], which may inspire the research on non-probabilistic quantitative rough set models.

The second direction is to build explanation-oriented approximations that emphasize on explaining and understanding the semantics. Most formulations of approximations focus on which objects should be included without due attention to their descriptions which are necessary to formulate rules. The lower and

upper approximations proposed by Pawlak in [18,19] and the three-way approximations proposed by Yao in [26] are defined as sets of objects, which are better illustrated by Fig. 1(b). A lack of internal structure leads to certain difficulties in deriving and interpreting classification rules. In contrast, a structured approximation [2,12] is defined as a set of definable sets instead of a set of objects. With clearer semantics, structured approximations can be conveniently and meaningfully applied to learn concepts with incomplete information, where most existing rough set models face a common challenge of interpreting their approximations in order to formulate classification rules [12,15].

This paper studies quantitative structured approximations as improvements in both directions. More specifically, we investigate quantitative structured approximations based on both probabilistic and non-probabilistic measures with both complete and incomplete information. This work focuses on exploring meaningful approaches to building explanation-oriented quantitative structured approximations. Accordingly, we present conceptual formulations of approximations that emphasize on the semantics, rather than computational formulations that emphasize efficient computations in practice. Further discussions on conceptual and computational formulations can be found in [4,12,16,25].

In the remainder of this paper, Sect. 2 reviews qualitative structured approximations with both complete (Sect. 2.1) and incomplete (Sect. 2.2) information. Section 3 explores the generalizations into quantitative structured approximations, including both complete (Sect. 3.1) and incomplete (Sects. 3.2 and 3.3) information. Conclusion and future work are discussed in Sect. 4.

2   Concept Analysis Using Qualitative Structured Three-Way Approximations

This section reviews the main results of qualitative structured approximations proposed in [12] with both complete and incomplete information.

2.1   Learning Intension with Complete Information

An information table is formally used in rough sets to represent a given dataset. In the case of complete information, a complete information table is formulated as the following tuple:

$$T = (OB, AT, \{V_a \mid a \in AT\}, \{I_a : OB \to V_a \mid a \in AT\}), \qquad (1)$$

where OB is a finite nonempty set of objects, AT is a finite nonempty set of attributes, V_a is the domain of an attribute a, and I_a is an information function which maps each object to a unique value in V_a. This unique value reflects the complete information or our complete knowledge. Logic formulas regarding attributes and their values are used as formal descriptions of objects. By arguing that a consideration of logic conjunction is sufficient for the rule-learning purpose, Hu and Yao [12] use a conjunctive description language DL_c consisting of formulas defined as follows:

(1) Atomic formulas: $(a = v) \in \mathrm{DL}_c$, where $a \in AT$ and $v \in V_a$;
(2) Composite formulas: if $p, q \in \mathrm{DL}_c$, and p and q do not share any attribute, then $p \wedge q \in \mathrm{DL}_c$.

The satisfiability of a formula by an object, denoted by $\models$, is defined as:

$$
\begin{aligned}
&(1)\ \ o \models (a = v) \Longleftrightarrow I_a(o) = v, \\
&(2)\ \ o \models (p \wedge q) \Longleftrightarrow o \models p \wedge o \models q,
\end{aligned}
\tag{2}
$$

where $o \in OB$, $a \in AT$, $v \in V_a$, and $p, q, p \wedge q \in \mathrm{DL}_c$. Accordingly, a formula is associated with a set of objects exhibiting its meaning.

Definition 1. *For a formula $p \in \mathrm{DL}_c$, the set of objects:*

$$
m(p) = \{o \in OB \mid o \models p\},
\tag{3}
$$

is called the meaning set of p.

On the other hand, objects in $m(p)$ can be uniformly described by p and thus, is considered to be definable. By using a formula as intension and its meaning set as extension, one can form a definable concept.

Definition 2. *A set of objects $O \subseteq OB$ is a conjunctively definable set if there exists a formula $p \in \mathrm{DL}_c$ such that $O = m(p)$. The pair $(p, m(p))$ is a conjunctively definable concept.*

$\mathrm{DEF}(T)$ is widely used to represent the family of definable sets in recent works [4, 22, 25]. Accordingly, we use $\mathrm{CDEF}(T)$ to represent the family of conjunctively definable concepts which is used to construct structured approximations.

Definition 3. *For a set of objects $X \subseteq OB$, its structured positive and negative regions [12] are defined as:*

$$
\begin{aligned}
\mathrm{SPOS}(X) &= \{(p, m(p)) \in \mathrm{CDEF}(T) \mid m(p) \neq \emptyset, m(p) \subseteq X\}, \\
\mathrm{SNEG}(X) &= \{(p, m(p)) \in \mathrm{CDEF}(T) \mid m(p) \neq \emptyset, m(p) \subseteq X^c\},
\end{aligned}
\tag{4}
$$

where X^c is the complement of X.

The boundary region is commonly defined through the positive and negative regions. With respect to Definition 3, the structured boundary region of X can be defined as:

$$
\mathrm{SBND}(X) = \{(p, m(p)) \in \mathrm{CDEF}(T) \mid \neg(m(p) \subseteq X) \wedge \neg(m(p) \subseteq X^c)\}.
\tag{5}
$$

From the view of learning intension, we are not interested in the boundary region since it doesn't lead to classification rules for recognizing either positive or negative instances of the concept.

Most research in the literature applies unstructured approximations [18, 19] which can be expressed as [12]:

$$
\begin{aligned}
\mathrm{POS}(X) &= \bigcup \{m(p) \mid (p, m(p)) \in \mathrm{SPOS}(X)\}, \\
\mathrm{NEG}(X) &= \bigcup \{m(p) \mid (p, m(p)) \in \mathrm{SNEG}(X)\}.
\end{aligned}
\tag{6}
$$

As argued and illustrated in [12], the structured approximations benefit rule learning with clear semantics obtained through preserving the internal structure as well as introducing intensions which are left-hand-sides of rules.

2.2 Learning Intension with Incomplete Information

Although an object actually takes exactly one value on an attribute, due to our limited or incomplete information, we may not be able to know this actual value. In such a case, an incomplete information table is used, which can be formally represented as the following tuple:

$$\widetilde{T} = (OB, AT, \{V_a \mid a \in AT\}, \{\widetilde{I}_a : OB \to 2^{V_a} - \{\emptyset\} \mid a \in AT\}), \qquad (7)$$

where OB, AT, and V_a have the same meanings as in a complete table, and $\widetilde{I}_a$ maps each object to a nonempty subset of V_a. Every value in $\widetilde{I}_a(x)$ may be the actual value of an object $x \in OB$ on an attribute $a \in AT$, but exactly one value is indeed the actual one which we do not know due to incomplete information.

Lipski [14] equivalently interprets an incomplete table as a family of complete tables. A complete table $(OB, AT, \{V_a \mid a \in AT\}, \{I_a : OB \to V_a \mid a \in AT\})$ is called a completion of $\widetilde{T}$ if:

$$\forall x \in OB \, \forall a \in AT (I_a(x) \in \widetilde{I}_a(x)). \qquad (8)$$

One gets a completion of $\widetilde{T}$ by picking up exactly one value for each object on each attribute. Since each value in $\widetilde{I}_a(x)$ represents one possibility of the actual value, a completion is a possibility of the actual table and called a possible world. Accordingly, Lipski's interpretation is called the possible-world semantics of an incomplete table. The family of all completions of $\widetilde{T}$ is denoted as $\text{COMP}(\widetilde{T})$.

The meaning set of a formula p in a completion T, denoted by $m(p|T)$, is a possibility of its actual meaning set. The collection of p's meaning sets in all completions covers all possibilities of p's actual meaning set and can be used to interpret p.

Definition 4. *The meaning set of a formula $p \in \text{DL}_c$ in an incomplete table $\widetilde{T}$ is defined as:*

$$\widetilde{m}(p) = \{m(p|T) \mid T \in \text{COMP}(\widetilde{T})\}, \qquad (9)$$

It is verified that $\widetilde{m}(p)$ is actually an interval set defined as [12]:

$$\widetilde{m}(p) = [m_*(p), m^*(p)] = \{S \subseteq OB \mid m_*(p) \subseteq S \subseteq m^*(p)\}. \qquad (10)$$

The sets $m_*(p)$ and $m^*(p)$ are the lower and upper bounds of $\widetilde{m}(p)$, respectively. The interval set $[m_*(p), m^*(p)]$ contains all sets in-between these two bounds (inclusive). Moreover, the two bounds can be computed as:

$$m_*(p) = \bigcap_{T \in \text{COMP}(\widetilde{T})} m(p|T) = \{x \in OB \mid \forall T \in \text{COMP}(\widetilde{T}), x \in m(p|T)\},$$

$$m^*(p) = \bigcup_{T \in \text{COMP}(\widetilde{T})} m(p|T) = \{x \in OB \mid \exists T \in \text{COMP}(\widetilde{T}), x \in m(p|T)\}. \quad (11)$$

The lower bound $m_*(p)$ contains objects satisfying p in every possible world, that is, they must satisfy p in the actual table and be included in p's actual meaning set. Similarly, the upper bound $m^*(p)$ contains objects satisfying p in at least one possible world, that is, they possibly satisfy p in the actual table and may be included in p's actual meaning set. By means of $\widetilde{m}(p)$, the definability can be generalized with respect to an incomplete table.

Definition 5. *An interval set $\mathcal{O}$ on OB is conjunctively definable if there exists a conjunctive formula $p \in \mathrm{DL}_c$ such that $\mathcal{O} = \widetilde{m}(p)$. The pair $(p, \widetilde{m}(p))$ is a conjunctively definable interval concept.*

The family of conjunctively definable interval concepts $\mathrm{CDEFI}(\widetilde{T})$ is used to construct the structured approximations in an incomplete table.

Definition 6. *Given a set of objects $X \subseteq OB$ in an incomplete table $\widetilde{T}$, two pairs of structured regions are constructed as [12]:*

(1) $\mathrm{SPOS}_*(X) = \{(p, \widetilde{m}(p)) \in \mathrm{CDEFI}(\widetilde{T}) \mid \forall S \in \widetilde{m}(p), S \neq \emptyset, S \subseteq X\}$,
$\mathrm{SNEG}_*(X) = \{(p, \widetilde{m}(p)) \in \mathrm{CDEFI}(\widetilde{T}) \mid \forall S \in \widetilde{m}(p), S \neq \emptyset, S \subseteq X^c\}$;

(2) $\mathrm{SPOS}^*(X) = \{(p, \widetilde{m}(p)) \in \mathrm{CDEFI}(\widetilde{T}) \mid \exists S \in \widetilde{m}(p), S \neq \emptyset, S \subseteq X\}$,
$\mathrm{SNEG}^*(X) = \{(p, \widetilde{m}(p)) \in \mathrm{CDEFI}(\widetilde{T}) \mid \exists S \in \widetilde{m}(p), S \neq \emptyset, S \subseteq X^c\}.$ (12)

$\mathrm{SPOS}_*(X)$ *and* $\mathrm{SNEG}_*(X)$ *are called lower structured regions, and* $\mathrm{SPOS}^*(X)$ *and* $\mathrm{SNEG}^*(X)$ *are upper structured regions.*

A lower structured region requires an exhaustivity of the set-inclusion relationship between a set in $\widetilde{m}(p)$ and X (or X^c), and an upper structured region requires an existence of such a relationship. The two lower structured regions give the lower bounds of the actual structured positive and negative regions, respectively, and the upper structured regions give the upper bounds [12].

3   Concept Analysis Using Quantitative Structured Three-Way Approximations

In this section, we generalize the qualitative structured regions into quantitative structured regions, in both complete and incomplete tables, based on two types of quantitative measures, namely, probabilities and subsethood measures. The generalization with respect to a complete table is straightforward based on existing research on quantitative unstructured approximations. In contrast, the generalization with respect to an incomplete table needs further investigation.

3.1   Probabilistic and Non-probabilistic Structured Approximations in a Complete Table

A probabilistic rough set model [27] replaces the qualitative set-inclusion with quantitative probabilities in defining unstructured approximations. By the same

idea, if an object described by $p \in \mathrm{DL}_c$ has a high probability of being a positive instance of X, then the concept $(p, m(p))$ is included in the structured positive region. This leads to the following probabilistic structured regions.

Definition 7. *For a set of objects $X \subseteq OB$, its probabilistic structured positive and negative regions are defined as:*

$$\mathrm{SPOS}^{pr}_{(\alpha, \cdot)}(X) = \{(p, m(p)) \in \mathrm{CDEF}(T) \mid Pr(X|m(p)) \geq \alpha\},$$
$$\mathrm{SNEG}^{pr}_{(\cdot, \gamma)}(X) = \{(p, m(p)) \in \mathrm{CDEF}(T) \mid Pr(X^c|m(p)) \geq \gamma\}, \qquad (13)$$

where $0 \leq \alpha, \gamma \leq 1$ are two thresholds, a dot represents a non-relevant threshold, and the probabilities are computed as:

$$Pr(X|m(p)) = \frac{|X \cap m(p)|}{|m(p)|}, \quad Pr(X^c|m(p)) = \frac{|X^c \cap m(p)|}{|m(p)|}. \qquad (14)$$

The qualitative structured regions can be viewed as a special case of the probabilistic structured regions with $\alpha = \gamma = 1$.
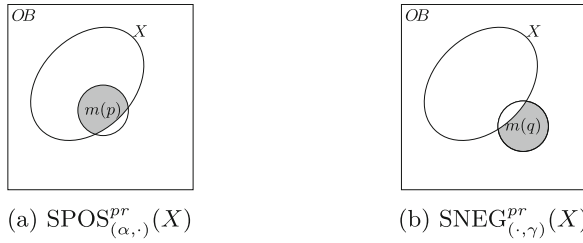

(a) $\mathrm{SPOS}^{pr}_{(\alpha, \cdot)}(X)$ (b) $\mathrm{SNEG}^{pr}_{(\cdot, \gamma)}(X)$

Fig. 2. Relationships between X and sets in $\mathrm{SPOS}^{pr}_{(\alpha, \cdot)}(X)$ and $\mathrm{SNEG}^{pr}_{(\cdot, \gamma)}(X)$

As shown in Fig. 2(a), $(p, m(p))$ is included in $\mathrm{SPOS}^{pr}_{(\alpha, \cdot)}(X)$ if the intersection between $m(p)$ and X occupies a large portion of $m(p)$. Accordingly, p is used to classify positive instances of X with an error rate:

$$\mathrm{IAE}((p, m(p))) = \frac{|X^c \cap m(p)|}{|m(p)|} = 1 - \frac{|X \cap m(p)|}{|m(p)|} \leq 1 - \alpha, \qquad (15)$$

which is called the rate of incorrect acceptance error (IAE) [5]. Similarly, a concept $(q, m(q)) \in \mathrm{SNEG}^{pr}_{(\cdot, \gamma)}(X)$ is associated with the following rate of incorrect rejection error (IRE):

$$\mathrm{IRE}((q, m(q))) = \frac{|X \cap m(q)|}{|m(q)|} = 1 - \frac{|X^c \cap m(q)|}{|m(q)|} \leq 1 - \gamma. \qquad (16)$$

At the expenses of IAE and IRE, we are able to approximate a larger part of X compared with the qualitative regions.

Most existing research on quantitative rough sets is based on the probabilistic (unstructured) approximations, such as decision-theoretic rough sets [29], game-theoretic rough sets [10], information-theoretic rough sets [6], naive Bayesian rough sets [30], confirmation-theoretic rough sets [9], and Bayesian rough sets [23]. Variable precision rough sets [33] can be viewed as both probabilistic and non-probabilistic in the sense that approximations are defined in terms of precisions which are estimated through probabilities.

In contrast, there is limited research [8,13,28] on non-probabilistic rough set models, which mainly uses subsethood measures and similarity measures instead of probabilities. Subsethood measure is a quantitative generalization of the qualitative set-inclusion. Given a universe OB, a normalized subsethood measure is defined as a mapping $sh : 2^{OB} \times 2^{OB} \to [0,1]$ where 2^{OB} is the power set of OB. For two sets $A, B \subseteq OB$, $sh(A \sqsubseteq B)$ represents the degree to which A is a subset of B. Yao and Deng [28] formulate quantitative unstructured approximations through subsethood measures, which unifies both probabilistic and non-probabilistic models. Following their formulation, we present the following quantitative structured regions based on subsethood measures.

Definition 8. *Suppose $sh : 2^{OB} \times 2^{OB} \to [0,1]$ is a normalized subsethood measure. For a given set of objects $X \subseteq OB$, its quantitative structured positive and negative regions can be defined as:*

$$\mathrm{SPOS}^{sh}_{(\alpha,\cdot)}(X) = \{(p, m(p)) \in \mathrm{CDEF}(T) \mid m(p) \neq \emptyset, sh(m(p) \sqsubseteq X) \geq \alpha\},$$

$$\mathrm{SNEG}^{sh}_{(\cdot,\gamma)}(X) = \{(p, m(p)) \in \mathrm{CDEF}(T) \mid m(p) \neq \emptyset, sh(m(p) \sqsubseteq X^c) \geq \gamma\}. (17)$$

By using a subsethood measure $sh(A \sqsubseteq B) = \frac{|A \cap B|}{|A|}$, we have:

$$sh(m(p) \sqsubseteq X) = \frac{|m(p) \cap X|}{|m(p)|} = Pr(X|m(p)), \qquad (18)$$

and similarly, $sh(m(p) \sqsubseteq X^c) = Pr(X^c|m(p))$. Thus, the two probabilistic regions in Definition 7 are special cases of the above two regions defined through subsethood measures. One may also consider many other meaningful cardinality-based subsethood measures [1] to formulate non-probabilistic approximations. For example, by using a measure $sh_{R_5^c}$ [11]:

$$sh_{R_5^c}(A \sqsubseteq B) = \begin{cases} \frac{|A^c|}{|(A \cap B)^c|}, & \neg(A = B = OB), \\ 1, & A = B = OB. \end{cases} \qquad (19)$$

we can formulate a pair of non-probabilistic structured regions of $X \subseteq OB$ as:

$$\mathrm{SPOS}^{sh_{R_5^c}}_{(\alpha,\cdot)}(X) = \{(p, m(p)) \in \mathrm{CDEF}(T) \mid m(p) \neq \emptyset, \frac{|m(p)^c|}{|(m(p) \cap X)^c|} \geq \alpha\},$$

$$\mathrm{SNEG}^{sh_{R_5^c}}_{(\cdot,\gamma)}(X) = \{(p, m(p)) \in \mathrm{CDEF}(T) \mid m(p) \neq \emptyset, \frac{|m(p)^c|}{|(m(p) \cap X^c)^c|} \geq \gamma\}. (20)$$

3.2   Probabilistic Structured Approximations in an Incomplete Table

One may also consider both probabilities and subsethood measures in defining quantitative regions in an incomplete table. In this subsection, we present two ways to define probabilistic structured regions. An intuitive way is to simply replace the set-inclusion in Definition 6 with probabilities, which leads to the following definition.

Definition 9. *Given a set of objects $X \subseteq OB$ in an incomplete table $\widetilde{T}$, one can construct the following probabilistic lower and upper structured regions:*

(1) $\mathrm{SPOS}^{pr}_{*(\alpha,\cdot)}(X) = \{(p, \widetilde{m}(p)) \mid \forall S \in \widetilde{m}(p), Pr(X|S) \geq \alpha\}$,

$\qquad \mathrm{SNEG}^{pr}_{*(\cdot,\gamma)}(X) = \{(p, \widetilde{m}(p)) \mid \forall S \in \widetilde{m}(p), Pr(X^c|S) \geq \gamma\}$;

(2) $\mathrm{SPOS}^{*pr}_{(\alpha,\cdot)}(X) = \{(p, \widetilde{m}(p)) \mid \exists S \in \widetilde{m}(p), Pr(X|S) \geq \alpha\}$,

$\qquad \mathrm{SNEG}^{*pr}_{(\cdot,\gamma)}(X) = \{(p, \widetilde{m}(p)) \mid \exists S \in \widetilde{m}(p), Pr(X^c|S) \geq \gamma\}$, $\qquad$ (21)

where the probabilities are computed as:

$$Pr(X|S) = \frac{|X \cap S|}{|S|}, \quad Pr(X^c|S) = \frac{|X^c \cap S|}{|S|}. \tag{22}$$

The condition $(p, \widetilde{m}(p)) \in \mathrm{CDEFI}(\widetilde{T})$ is omitted in Definition 9 and the following definitions where this doesn't cause misunderstanding. Since $\widetilde{m}(p)$ is a collection of p's meaning sets in all completions, the above regions can be equivalently expressed through the family $\mathrm{COMP}(\widetilde{T})$.

Proposition 1. *Given a set of objects $X \subseteq OB$, one may construct the probabilistic lower and upper structured regions of X as follows:*

(1) $\mathrm{SPOS}^{pr}_{*(\alpha,\cdot)}(X) = \{(p, \widetilde{m}(p)) \mid \forall T \in \mathrm{COMP}(\widetilde{T}), Pr(X|m(p|T)) \geq \alpha\}$,

$\qquad \mathrm{SNEG}^{pr}_{*(\cdot,\gamma)}(X) = \{(p, \widetilde{m}(p)) \mid \forall T \in \mathrm{COMP}(\widetilde{T}), Pr(X^c|m(p|T)) \geq \gamma\}$;

(2) $\mathrm{SPOS}^{*pr}_{(\alpha,\cdot)}(X) = \{(p, \widetilde{m}(p)) \mid \exists T \in \mathrm{COMP}(\widetilde{T}), Pr(X|m(p|T)) \geq \alpha\}$,

$\qquad \mathrm{SNEG}^{*pr}_{(\cdot,\gamma)}(X) = \{(p, \widetilde{m}(p)) \mid \exists T \in \mathrm{COMP}(\widetilde{T}), Pr(X^c|m(p|T)) \geq \gamma\}$. (23)

The possible-world semantics also connects the above probabilistic regions in an incomplete table to those in a complete table (i.e., Definition 7). The condition $Pr(X|m(p|T)) \geq \alpha$ implies that the concept $(p, m(p|T))$ is included in $\mathrm{SPOS}^{pr}_{(\alpha,\cdot)}(X|T)$. Accordingly, we get the following theorem.

Theorem 1. *Given a set of objects $X \subseteq OB$, we have:*

(1) $(p, \widetilde{m}(p)) \in \mathrm{SPOS}^{pr}_{*(\alpha,\cdot)}(X) \Leftrightarrow \forall T \in \mathrm{COMP}(\widetilde{T}), (p, m(p|T)) \in \mathrm{SPOS}^{pr}_{(\alpha,\cdot)}(X|T)$,

$\qquad (p, \widetilde{m}(p)) \in \mathrm{SNEG}^{pr}_{*(\cdot,\gamma)}(X) \Leftrightarrow \forall T \in \mathrm{COMP}(\widetilde{T}), (p, m(p|T)) \in \mathrm{SNEG}^{pr}_{(\cdot,\gamma)}(X|T)$;

(2) $(p, \widetilde{m}(p)) \in \mathrm{SPOS}^{*pr}_{(\alpha,\cdot)}(X) \Leftrightarrow \exists T \in \mathrm{COMP}(\widetilde{T}), (p, m(p|T)) \in \mathrm{SPOS}^{pr}_{(\alpha,\cdot)}(X|T)$,

$\qquad (p, \widetilde{m}(p)) \in \mathrm{SNEG}^{*pr}_{(\cdot,\gamma)}(X) \Leftrightarrow \exists T \in \mathrm{COMP}(\widetilde{T}), (p, m(p|T)) \in \mathrm{SNEG}^{pr}_{(\cdot,\gamma)}(X|T)$.

$$\tag{24}$$

Definition 9 and Proposition 1 provide two mathematically equivalent but semantically different formulations. Definition 9 is a straightforward generalization of the qualitative regions. Proposition 1 provides an equivalent version through the family $\mathrm{COMP}(\widetilde{T})$, which offers a clearer semantics. This clear semantics enables us to explore the relationships stated in Theorem 1.

Instead of considering $Pr(X|S)$ for every set $S \in \widetilde{m}(p)$, we may generalize $Pr(X|S)$ into a probability $Pr(X|\widetilde{m}(p))$ regarding a set X and an interval set $\widetilde{m}(p)$ which has not been well studied. Different interpretations of $Pr(X|\widetilde{m}(p))$ may lead to different formulas. In our work, we interpret $Pr(X|\widetilde{m}(p))$ as the probability of a set in $\widetilde{m}(p)$ being included in X. Accordingly, we define the following probabilistic structured regions.

Definition 10. *Given a set of objects $X \subseteq OB$, one may define the following pair of probabilistic structured positive and negative regions:*

$$\widetilde{\mathrm{SPOS}}^{pr}_{(\alpha,\cdot)}(X) = \{(p, \widetilde{m}(p)) \in \mathrm{CDEFI}(\widetilde{T}) \mid Pr(X|\widetilde{m}(p)) \geq \alpha\},$$
$$\widetilde{\mathrm{SNEG}}^{pr}_{(\cdot,\gamma)}(X) = \{(p, \widetilde{m}(p)) \in \mathrm{CDEFI}(\widetilde{T}) \mid Pr(X^c|\widetilde{m}(p)) \geq \gamma\}, \qquad (25)$$

where the probabilities are computed as:

$$Pr(X|\widetilde{m}(p)) = \frac{|\{S \in \widetilde{m}(p)|\emptyset \neq S \subseteq X\}|}{|\widetilde{m}(p)|},$$
$$Pr(X^c|\widetilde{m}(p)) = \frac{|\{S \in \widetilde{m}(p)|\emptyset \neq S \subseteq X^c\}|}{|\widetilde{m}(p)|}. \qquad (26)$$

Since each set in $\widetilde{m}(p)$ represents a possibility of p's actual meaning set, a high probability $Pr(X|\widetilde{m}(p))$ means that, in a large portion of all possible worlds $\mathrm{COMP}(\widetilde{T})$, the meaning set of p is included in X, or equivalently, p appears in the qualitative structured positive region of X. Thus, it is with high probability that p appears in the actual qualitative structured positive region of X.

Proposition 2. *Given a set of objects $X \subseteq OB$, we have:*

$$(p, \widetilde{m}(p)) \in \widetilde{\mathrm{SPOS}}^{pr}_{(\alpha,\cdot)}(X) \Longrightarrow Pr\big((p, m(p|T_0)) \in \mathrm{SPOS}(X|T_0)\big) \geq \alpha,$$
$$(p, \widetilde{m}(p)) \in \widetilde{\mathrm{SNEG}}^{pr}_{(\cdot,\gamma)}(X) \Longrightarrow Pr\big((p, m(p|T_0)) \in \mathrm{SNEG}(X|T_0)\big) \geq \gamma, \quad (27)$$

where $T_0 \in \mathrm{COMP}(\widetilde{T})$ is the actual table.

The two probabilities $Pr(X|\widetilde{m}(p))$ and $Pr(X^c|\widetilde{m}(p))$ can be efficiently computed through the two bounds of $\widetilde{m}(p)$. For $Pr(X|\widetilde{m}(p))$, if $m_*(p) \not\subseteq X$, then no set in $\widetilde{m}(p)$ is included X, that is, $Pr(X|\widetilde{m}(p)) = 0$. Otherwise, we have:

$$
\begin{aligned}
Pr(X|\widetilde{m}(p)) &= \frac{|\{S \subseteq OB \mid m_*(p) \subseteq S \subseteq m^*(p) \wedge S \subseteq X\}|}{|\{S \subseteq OB \mid m_*(p) \subseteq S \subseteq m^*(p)\}|} \\
&= \frac{|\{S \subseteq OB \mid m_*(p) \subseteq S \subseteq m^*(p) \cap X\}|}{|\{S \subseteq OB \mid m_*(p) \subseteq S \subseteq m^*(p)\}|} = \frac{2^{|m^*(p) \cap X - m_*(p)|}}{2^{|m^*(p) - m_*(p)|}} \\
&= \frac{2^{|m^*(p) \cap X| - |m_*(p)|}}{2^{|m^*(p)| - |m_*(p)|}} = 2^{|m^*(p) \cap X| - |m^*(p)|}. \qquad (28)
\end{aligned}
$$

The probability $Pr(X^c|\widetilde{m}(p))$ can be similarly computed and the following computational formulation of the structured regions can be accordingly obtained.

Theorem 2. *Given a set of objects $X \subseteq OB$, one may construct a pair of probabilistic structured regions of X as:*

$$\widetilde{\mathrm{SPOS}}^{pr}_{(\alpha,\cdot)}(X) = \{(p,\widetilde{m}(p)) \in \mathrm{CDEFI}(\widetilde{T}) \mid \mu \cdot 2^{|m^*(p)\cap X| - |m^*(p)|} \geq \alpha\},$$

$$\widetilde{\mathrm{SNEG}}^{pr}_{(\cdot,\gamma)}(X) = \{(p,\widetilde{m}(p)) \in \mathrm{CDEFI}(\widetilde{T}) \mid \mu_c \cdot 2^{|m^*(p)\cap X^c| - |m^*(p)|} \geq \gamma\}, \quad (29)$$

where μ and μ_c are two numbers defined as:

$$\mu = \begin{cases} 1, & \text{if } m_*(p) \subseteq X, \\ 0, & \text{otherwise.} \end{cases} \qquad \mu_c = \begin{cases} 1, & \text{if } m_*(p) \subseteq X^c, \\ 0, & \text{otherwise.} \end{cases} \quad (30)$$

While Definition 10 provides a conceptual understanding of the structured regions which requires an exhaustive scan of $\widetilde{m}(p)$ to compute the probabilities, Theorem 2 gives an equivalent computational formulation where the probabilities can be efficiently computed through the two bounds of $\widetilde{m}(p)$.

Example 1. We illustrate the above probabilistic structured regions in Definitions 9 and 10 with an incomplete table given by Table 1 [12]. The family $\mathrm{CDEFI}(\widetilde{T})$ is given by Table 2.

Table 1. An incomplete table $\widetilde{T}$ [12]

	a_1	a_2	a_3
o_1	{1}	{5}	{6}
o_2	{2}	{4}	{6}
o_3	{1}	{3}	{6}
o_4	{1}	{3,4}	{7}
o_5	{1,2}	{5}	{6}
o_6	{1}	{4}	{6}
o_7	{1,2}	{4}	{6}

Given a set of objects $X = \{o_2, o_5, o_6, o_7\}$ ($X^c = \{o_1, o_3, o_4\}$) and thresholds $\alpha = \gamma = 0.7$, the two pairs of lower and upper regions defined in Definition 9 are:

$$\begin{aligned}
\mathrm{SPOS}^{pr}_{*(\alpha,\cdot)}(X) =\ & \{IC_2, IC_4, IC_{12}, IC_{16}, IC_{20}, IC_{26}, IC_{32}\} \\
\mathrm{SNEG}^{pr}_{*(\cdot,\gamma)}(X) =\ & \{IC_3, IC_7, IC_8, IC_{15}, IC_{18}, IC_{24}\} \\
\mathrm{SPOS}^{*pr}_{(\alpha,\cdot)}(X) =\ & \{IC_2, IC_4, IC_9, IC_{12}, IC_{13}, IC_{16}, IC_{20}, IC_{26}, IC_{32}, IC_{34}\} \\
\mathrm{SNEG}^{*pr}_{(\cdot,\gamma)}(X) =\ & \{IC_1, IC_3, IC_7, IC_8, IC_{10}, IC_{15}, IC_{18}, IC_{19}, IC_{21}, IC_{24}, IC_{25}, \\
& IC_{27}, IC_{28}\}. \quad (31)
\end{aligned}$$

Table 2. The family CDEFI($\widetilde{T}$) for Table 1

Label	Intension	Extension	Label	Intension	Extension
IC_1	$a_1 = 1$	$[\{o_1, o_3, o_4, o_6\},$ $\{o_1, o_3, o_4, o_5, o_6, o_7\}]$	IC_{19}	$a_2 = 3 \wedge a_3 = 7$	$[\emptyset, \{o_4\}]$
IC_2	$a_1 = 2$	$[\{o_2\}, \{o_2, o_5, o_7\}]$	IC_{20}	$a_2 = 4 \wedge a_3 = 6$	$[\{o_2, o_6, o_7\}, \{o_2, o_6, o_7\}]$
IC_3	$a_2 = 3$	$[\{o_3\}, \{o_3, o_4\}]$	IC_{21}	$a_2 = 4 \wedge a_3 = 7$	$[\emptyset, \{o_4\}]$
IC_4	$a_2 = 4$	$[\{o_2, o_6, o_7\},$ $\{o_2, o_4, o_6, o_7\}]$	IC_{22}	$a_2 = 5 \wedge a_3 = 6$	$[\{o_1, o_5\}, \{o_1, o_5\}]$
IC_5	$a_2 = 5$	$[\{o_1, o_5\}, \{o_1, o_5\}]$	IC_{23}	$a_2 = 5 \wedge a_3 = 7$	$[\emptyset, \emptyset]$
IC_6	$a_3 = 6$	$[\{o_1, o_2, o_3, o_5, o_6, o_7\},$ $\{o_1, o_2, o_3, o_5, o_6, o_7\}]$	IC_{24}	$a_1 = 1 \wedge a_2 = 3 \wedge a_3 = 6$	$[\{o_3\}, \{o_3\}]$
IC_7	$a_3 = 7$	$[\{o_4\}, \{o_4\}]$	IC_{25}	$a_1 = 1 \wedge a_2 = 3 \wedge a_3 = 7$	$[\emptyset, \{o_4\}]$
IC_8	$a_1 = 1 \wedge a_2 = 3$	$[\{o_3\}, \{o_3, o_4\}]$	IC_{26}	$a_1 = 1 \wedge a_2 = 4 \wedge a_3 = 6$	$[\{o_6\}, \{o_6, o_7\}]$
IC_9	$a_1 = 1 \wedge a_2 = 4$	$[\{o_6\}, \{o_4, o_6, o_7\}]$	IC_{27}	$a_1 = 1 \wedge a_2 = 4 \wedge a_3 = 7$	$[\emptyset, \{o_4\}]$
IC_{10}	$a_1 = 1 \wedge a_2 = 5$	$[\{o_1\}, \{o_1, o_5\}]$	IC_{28}	$a_1 = 1 \wedge a_2 = 5 \wedge a_3 = 6$	$[\{o_1\}, \{o_1, o_5\}]$
IC_{11}	$a_1 = 2 \wedge a_2 = 3$	$[\emptyset, \emptyset]$	IC_{29}	$a_1 = 1 \wedge a_2 = 5 \wedge a_3 = 7$	$[\emptyset, \emptyset]$
IC_{12}	$a_1 = 2 \wedge a_2 = 4$	$[\{o_2\}, \{o_2, o_7\}]$	IC_{30}	$a_1 = 2 \wedge a_2 = 3 \wedge a_3 = 6$	$[\emptyset, \emptyset]$
IC_{13}	$a_1 = 2 \wedge a_2 = 5$	$[\emptyset, \{o_5\}]$	IC_{31}	$a_1 = 2 \wedge a_2 = 3 \wedge a_3 = 7$	$[\emptyset, \emptyset]$
IC_{14}	$a_1 = 1 \wedge a_3 = 6$	$[\{o_1, o_3, o_6\},$ $\{o_1, o_3, o_5, o_6, o_7\}]$	IC_{32}	$a_1 = 2 \wedge a_2 = 4 \wedge a_3 = 6$	$[\{o_2\}, \{o_2, o_7\}]$
IC_{15}	$a_1 = 1 \wedge a_3 = 7$	$[\{o_4\}, \{o_4\}]$	IC_{33}	$a_1 = 2 \wedge a_2 = 4 \wedge a_3 = 7$	$[\emptyset, \emptyset]$
IC_{16}	$a_1 = 2 \wedge a_3 = 6$	$[\{o_2\}, \{o_2, o_5, o_7\}]$	IC_{34}	$a_1 = 2 \wedge a_2 = 5 \wedge a_3 = 6$	$[\emptyset, \{o_5\}]$
IC_{17}	$a_1 = 2 \wedge a_3 = 7$	$[\emptyset, \emptyset]$	IC_{35}	$a_1 = 2 \wedge a_2 = 5 \wedge a_3 = 7$	$[\emptyset, \emptyset]$
IC_{18}	$a_2 = 3 \wedge a_3 = 6$	$[\{o_3\}, \{o_3\}]$			

With the same set X and the same thresholds, the pair of regions defined in Definition 10 are:

$$\widetilde{\mathrm{SPOS}}^{pr}_{(\alpha,\cdot)}(X) = \{IC_2, IC_{12}, IC_{16}, IC_{20}, IC_{26}, IC_{32}\}$$
$$\widetilde{\mathrm{SNEG}}^{pr}_{(\cdot,\gamma)}(X) = \{IC_3, IC_7, IC_8, IC_{15}, IC_{18}, IC_{24}\}. \tag{32}$$

3.3   Non-probabilistic Structured Approximations in an Incomplete Table Based on Subsethood Measures

We present a more general formulation of quantitative structured regions by using a subsethood measure instead of the probabilities in Definition 9.

Definition 11. *Suppose* $sh : 2^{OB} \times 2^{OB} \rightarrow [0,1]$ *is a normalized subsethood measure. Given a set* $X \subseteq OB$, *one can define the following structured regions:*

(1) $\mathrm{SPOS}^{sh}_{*(\alpha,\cdot)}(X) = \{(p, \widetilde{m}(p)) \mid \forall S \in \widetilde{m}(p), S \neq \emptyset, sh(S \sqsubseteq X) \geq \alpha\},$

   $\mathrm{SNEG}^{sh}_{*(\cdot,\gamma)}(X) = \{(p, \widetilde{m}(p)) \mid \forall S \in \widetilde{m}(p), S \neq \emptyset, sh(S \sqsubseteq X^c) \geq \gamma\};$

(2) $\mathrm{SPOS}^{*sh}_{(\alpha,\cdot)}(X) = \{(p, \widetilde{m}(p)) \mid \exists S \in \widetilde{m}(p), S \neq \emptyset, sh(S \sqsubseteq X) \geq \alpha\},$

   $\mathrm{SNEG}^{*sh}_{(\cdot,\gamma)}(X) = \{(p, \widetilde{m}(p)) \mid \exists S \in \widetilde{m}(p), S \neq \emptyset, sh(S \sqsubseteq X^c) \geq \gamma\}. \tag{33}$

By using a subsethood measure $sh(A \sqsubseteq B) = \frac{|A \cap B|}{|A|}$, the probabilistic regions in Definition 9 become special cases of the above regions. Non-probabilistic struc-

tured regions can be constructed by applying non-probabilistic subsethood measures such as $sh_{R_5^c}$ given in Eq. (19). Since each set $S \in \widetilde{m}(p)$ is a meaning set of p in a completion, one can equivalently express the above regions through the family $\text{COMP}(\widetilde{T})$, for example:

$$\text{SPOS}^{sh}_{*(\alpha,\cdot)}(X) = \{(p, \widetilde{m}(p)) \mid \forall\, T \in \text{COMP}(\widetilde{T}), m(p|T) \neq \emptyset, sh(m(p|T) \sqsubseteq X) \geq \alpha\}. \quad (34)$$

Similar as in Theorem 1, one may also establish relationships between the above subsethood-based regions and those in the completions, for example:

$$(p, \widetilde{m}(p)) \in \text{SPOS}^{sh}_{*(\alpha,\cdot)}(X) \Longleftrightarrow \forall\, T \in \text{COMP}(\widetilde{T}), (p, m(p|T)) \in \text{SPOS}^{sh}_{(\alpha,\cdot)}(X|T). \quad (35)$$

Alternatively, one may generalize subsethood measures for two sets into those for an interval set and a set to construct quantitative regions. Such subsethood measures evaluate the degree to which an interval set is included in a set.

Definition 12. *Suppose* $\mathcal{S}h : \mathcal{I}(OB) \times 2^{OB} \to [0,1]$ *is a normalized subsethood measure where* $\mathcal{I}(OB)$ *is the family of interval sets on* OB. *Given a set of objects* $X \subseteq OB$, *one may define the following pair of structured regions:*

$$\widetilde{\text{SPOS}}^{\mathcal{S}h}_{(\alpha,\cdot)}(X) = \{(p, \widetilde{m}(p)) \in \text{CDEFI}(\widetilde{T}) \mid \mathcal{S}h(\widetilde{m}(p) \sqsubseteq X) \geq \alpha\},$$

$$\widetilde{\text{SNEG}}^{\mathcal{S}h}_{(\cdot,\gamma)}(X) = \{(p, \widetilde{m}(p)) \in \text{CDEFI}(\widetilde{T}) \mid \mathcal{S}h(\widetilde{m}(p) \sqsubseteq X^c) \geq \gamma\}. \quad (36)$$

This definition depends on the specific definition of $\mathcal{S}h$ which has not been well studied. One may define $\mathcal{S}h(\mathcal{A} \sqsubseteq B)$ through $sh(A \sqsubseteq B)$ where $A \in \mathcal{A}$, such as taking the average:

$$\mathcal{S}h(\mathcal{A} \sqsubseteq B) = \frac{\sum\limits_{A \in \mathcal{A}} sh(A \sqsubseteq B)}{|\mathcal{A}|}. \quad (37)$$

One may also define $\mathcal{S}h$ through the qualitative set-inclusion such as using the proportion of subsets of B in $\mathcal{A}$:

$$\mathcal{S}h(\mathcal{A} \sqsubseteq B) = \frac{|\{A \in \mathcal{A} \mid \emptyset \neq A \subseteq B\}|}{|\mathcal{A}|}. \quad (38)$$

With the latter, the probabilistic regions in Definition 10 become special cases of the above regions in Definition 12. One may also construct non-probabilistic quantitative regions through Definition 12 by applying a subsethood measure $\mathcal{S}h$ that cannot be explained through probabilities.

4   Conclusion and Future Work

To combine the advantages of both quantitative and structured approximations, this paper investigates quantitative formulations of structured approximations in both complete and incomplete tables. We consider both probabilistic formulations which are widely studied in the literature regarding unstructured

approximations and non-probabilistic formulations which have not received due attention.

Our work brings up several interesting topics to work on. A first topic is the interpretation and determination of thresholds in various quantitative regions. While there are lots of existing related studies with respect to the probabilistic unstructured approximations in a complete table, the thresholds in subsethood-based quantitative regions and those regions in incomplete tables need further investigation. Solutions to this topic will help construct efficient computational formulations of approximations, which is a second topic for future work. A third topic is to investigate the relationships between this work and other concept analysis approaches such as lattice theory, formal concept analysis, and pattern structures. A fourth topic is the generalization of subsethood measures $sh(A \sqsubseteq B)$ regarding two sets into those regarding interval sets, including $\mathcal{S}h(\mathcal{A} \sqsubseteq B)$, $s\mathcal{H}(A \sqsubseteq \mathcal{B})$, and $\mathcal{S}\mathcal{H}(\mathcal{A} \sqsubseteq \mathcal{B})$ where $\mathcal{A}, \mathcal{B}$ are interval sets and A, B are sets. The research on this topic will shed new light on defining meaningful quantitative approximations in an incomplete table.

References

1. De Baets, B., De Meyer, H., Naessens, H.: On rational cardinality-based inclusion measures. Fuzzy Sets Syst. **128**, 169–183 (2002)
2. Bryniarski, E.: A calculus of rough sets of the first order. Bull. Polish Acad. Sci. Math. **37**, 71–78 (1989)
3. Buroker, J.: Port Royal Logic. Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/port-royal-logic/. Accessed January 2020
4. D'eer, L., Cornelis, C., Yao, Y.: A semantically sound approach to Pawlak rough sets and covering-based rough sets. Int. J. Approximate Reasoning **78**, 62–72 (2016)
5. Deng, X.: Three-way classification models. PhD thesis, University of Regina (2015)
6. Deng, X., Yao, Y.: A multifaceted analysis of probabilistic three-way decisions. Fundamenta Informaticae **132**, 291–313 (2014)
7. Ganter, B., Stumme, G., Wille, R. (eds.): Formal Concept Analysis: Foundations and Applications. LNCS (LNAI), vol. 3626. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-31881-1
8. Gomolińska, A.: Rough approximation based on weak q-RIFs. In: Peters, J.F., Skowron, A., Wolski, M., Chakraborty, M.K., Wu, W.-Z. (eds.) Transactions on Rough Sets X. LNCS, vol. 5656, pp. 117–135. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-03281-3_4
9. Greco, S., Matarazzo, B., Slowinski, R.: Parameterized rough set model using rough membership and Bayesian confirmation measures. Int. J. Approximate Reasoning **49**, 285–300 (2008)
10. Herbert, J.P., Yao, J.T.: Game-theoretic rough sets. Fundamenta Informaticae **108**, 267–286 (2011)
11. Hu, M., Deng, X., Yao, Y.: On the properties of subsethood measures. Inf. Sci. **494**, 208–232 (2019)

12. Hu, M., Yao, Y.: Structured approximations as a basis for three-way decisions with rough sets. Knowl.-Based Syst. **165**, 92–109 (2019)
13. Janicki, R., Lenarcic, A.: Optimal approximations with rough sets and similarities in measure spaces. Int. J. Approximate Reasoning **71**, 1–14 (2016)
14. Lipski, W.: On semantics issues connected with incomplete information table. ACM Trans. Database Syst. **4**, 262–296 (1979)
15. Luo, J., Fujita, H., Yao, Y., Qin, K.: On modeling similarity and three-way decision under incomplete information in rough set theory. Knowl.-Based Syst. **191**, 105251 (2020)
16. Luo, J., Hu, M., Qin, K.: Three-way decision with incomplete information based on similarity and satisfiability. Int. J. Approximate Reasoning **120**, 151–183 (2020)
17. Naidenova, X., Buzmakov, A., Parkhomenko, V., Schukin, A.: Notes on relation between symbolic classifiers. In: Watson, B.W., Kuznetsov, S.O. (eds.) CEUR Workshop Proceedings, CEUR-WS, vol. 1921, pp. 88–103 (2017)
18. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic, Dordrecht (1991)
19. Pawlak, Z.: Rough sets. Int. J. Comput. Inf. Sci. **11**, 341–356 (1982)
20. Qi, J., Qian, T., Wei, L.: The connections between three-way and classical concept lattices. Knowl.-Based Syst. **91**, 143–151 (2016)
21. Ren, R., Wei, L., Yao, Y.: An analysis of three types of partially-known formal concepts. Int. J. Mach. Learn. Cybernet. **9**(11), 1767–1783 (2017). https://doi.org/10.1007/s13042-017-0743-z
22. Sang, B., Yang, L., Chen, H., Xu, W., Guo, Y., Yuan, Z.: Generalized multigranulation double-quantitative decision-theoretic rough set of multi-source information system. Int. J. Approximate Reasoning **115**, 157–179 (2019)
23. Slezak, D., Ziarko, W.: Bayesian rough set model. In: Proceedings of the International Workshop on Foundation of Data Mining, pp. 131–135 (2002)
24. Yao, Y.: Three-way granular computing, rough sets, and formal concept analysis. Int. J. Approximate Reasoning **116**, 106–125 (2020)
25. Yao, Y.: The two sides of the theory of rough sets. Knowl.-Based Syst. **80**, 67–77 (2015)
26. Yao, Y.: Three-way decision: an interpretation of rules in rough set theory. In: Wen, P., Li, Y., Polkowski, L., Yao, Y., Tsumoto, S., Wang, G. (eds.) RSKT 2009. LNCS (LNAI), vol. 5589, pp. 642–649. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-02962-2_81
27. Yao, Y.: Probabilistic rough set approximations. Int. J. Approximate Reasoning **49**, 255–271 (2008)
28. Yao, Y., Deng, X.: Quantitative rough sets based on subsethood measures. Inf. Sci. **267**, 306–322 (2014)
29. Yao, Y., Wong, S.K.M., Lingras, P.J.: A decision-theoretic rough set model. In: Proceedings of the 5th International Symposium on Methodologies for Intelligent Systems, vol. 5, pp. 17–24 (1990)
30. Yao, Y., Zhou, B.: Naive Bayesian rough sets. In: Yu, J., Greco, S., Lingras, P., Wang, G., Skowron, A. (eds.) RSKT 2010. LNCS (LNAI), vol. 6401, pp. 719–726. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-16248-0_97
31. Zhang, T., Li, H., Liu, M., Rong, M.: Incremental concept-cognitive learning based on attribute topology. Int. J. Approximate Reasoning **118**, 173–189 (2020)
32. Zhi, H., Qi, J., Qian, T., Wei, L.: Three-way dual concept analysis. Int. J. Approximate Reasoning **114**, 151–165 (2020)
33. Ziarko, W.: Variable precision rough set model. J. Comput. Syst. Sci. **46**, 39–59 (1993)