Annals of **F**uzzy **M**athematics and **I**nformatics
Volume x, No. x, (mm 201y), pp. 1–xx
ISSN: 2093–9310 (print version)
ISSN: 2287–6235 (electronic version)
http://www.afmi.or.kr

@FMI

Optimal approximations of rough sets

Rui Qin, Zhi ZhengXiao, Keyun Qin

Received 28 July 2025; Revised 5 December 2025

Abstract. In this study, the optimal approximations of rough sets in an approximation space is proposed based on similarity measures induced by symmetric difference. Some basic properties of optimal approximations are investigated. The variation rules of similarity degrees between the target concepts and their optimal approximations in different granularity spaces are surveyed. The attribute reduction approaches for decision systems which preserve the optimal approximations of decision classes unchanged are presented by using discernibility functions. In this case, the attribute reductions can be computed all at once.

2020 AMS Classification: 03E72, 08A72

Keywords: Rough set, Similarity measure, Optimal approximations, Decision system, Attribute reduction.

1. Introduction

The theory of rough sets was firstly proposed by Pawlak [1]. It is an extension of set theory and provides a systematic method for dealing with vague concepts caused by indiscernibility in situations with insufficient and incomplete information. Using the notions of upper and lower approximations in rough set theory, knowledge hidden in information systems may be unravelled and expressed in the form of decision rules [2, 3, 4, 5, 6]. Also, some extended rough set models, such as general binary relation based rough set [7], fuzzy rough set [8], variable precision rough set [9], probabilistic rough set [10] and so on were presented to extend the application scope of rough sets.

In rough set model, the upper and lower approximations are two boundary lines of the target set. The lower approximation of a set is the largest definable set contained in it and the upper approximation of a set is the least definable set containing it. In real application, whether the lower or upper approximation is to be used is not definite. There is little discussion in original rough set theory on how to establish one crisp set as approximation set of a target set instead of only giving out two boundary lines. Considering this issue, Janicki and Lenarci [11] presented a metric approach to rough sets. The notion of optimal approximation in rough approximation spaces with respect to a measure of similarity is proposed, and an algorithm to compute optimal approximation using the Jaccard similarity measure is surveyed. The optimal approximation may be closer to the target set compared with upper approximation and lower approximation. In [12], Janicki and Lenarci proposed finite and null-free measure on the universe to generate the similarity measures. The algorithm is presented to compute the optimal approximation by using the Marczewski-Steinhaus similarity measure. The algorithm is also effective for similarity measures which are consistent with the Marczewski-Steinhaus similarity measure. Based on Jaccard similarity measure, Zhang et al [13] proposed a method for building an approximation set of the uncertain set X. It is proved that $R_{0.5}(X)$ is the optimal approximation set of X in some cases. The changing regularities of similarity between $R_{0.5}(X)$ and X with the change of knowledge granularity in knowledge space are surveyed. Zhang et al [14, 15] proposed the notion of relative approximate degree in the decision making information systems which can well reflect the degree of similarity between the division of domain made by condition and decision attributes. A heuristic attribute reduction method and algorithm based on approximation set of rough set is proposed. In [16], a kind of fuzzy similarity between a target set and its approximation set is presented based on Euclidean distance between two fuzzy sets. It is proved that 0.5-approximation set is the best approximation set of a rough set in all definable sets. The change laws of fuzzy similarity between a target set and its 0.5-approximation set in different granularity spaces is analyzed in detail.

We note that there are many kinds of similarity measures in set theory for cooping with different kinds of application problems. In some cases, the similarity degree of two sets not only related to these two sets, but also related to the cardinality of the whole universe. Jaccard similarity measure can not well reflect this characteristics. The existing studies on optimal approximation rough sets are mainly based on Jaccard similarity measure. In this study, we consider a new similarity measure in the study of optimal approximation sets and investigate related attribute reduction problems.

2. Preliminaries

In this section, we briefly introduce basic notions related to rough sets. We refer to [1] for details.

Definition 2.1 ([1])**.** Let U be a nonempty set and R an equivalence relation on U. (U, R) is called an *approximation space*. For any $X \subseteq U$, the sets

$$\underline{R}(X) = \{x | [x]_R \subseteq X\} \tag{2.1}$$

and

$$\overline{R}(X) = \{x | [x]_R \cap X \neq \varnothing\} \tag{2.2}$$

are called the *lower* and *upper approximations* of X respectively, where $[x]_R = \{y \in U | (x, y) \in R\}$ is the equivalence class with respect to R which containing x.

X is called a R *definable set*, if $\underline{R}(X) = \overline{R}(X)$. That is to say, X is a R definable set if and only if X is the union of some R equivalence classes. A definable set Y is called the *optimal approximation* of X [11], if $S(X,Y) \geq S(X,Z)$ for any definable set Z, where S is a similarity measure. Based on Jaccard similarity measure, some basic properties of optimal approximation are surveyed.

Theorem 2.2 ([13])**.** *Let (U,R) be an approximation space, $X \subseteq U$ and $R_{0.5}(X) = \{x \in U | \frac{|[x]_R \cap X|}{|[x]_R|} \geq 0.5\}$.*

(1) $S_J(X, R_{0.5}(X)) \geq S_J(X, \underline{R}(X))$,

(2) if $\frac{|X|}{\overline{R}(X)} > \frac{|X - R_{0.5}(X)|}{|\overline{R}(X) - R_{0.5}(X) - X|}$, then $S_J(X, R_{0.5}(X)) \geq S_J(X, \overline{R}(X))$,

where $S_J(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}$ is the Jaccard similarity measure.

Algorithm 1 (Finding the Greatest Optimal Approximation)[11]

Let (U,R) be an approximation space, $X \subseteq U$.

(1) Construct $\underline{R}(X)$, $\overline{R}(X)$ and $B(X) = \{[x]_R | [x]_R \nsubseteq X \wedge [x]_R \cap X \neq \varnothing\}$. Assume $r = |B(X)|$.

(2) For each $[x]_R \in B(X)$, calculate $\alpha([x]_R) = \frac{|[x]_R \cap X|}{|[x]_R - X|}$.

(3) Order $\alpha([x]_R)$ in decreasing order and number the elements of $B(X)$ by this order, so $B(X) = \{[x_1]_R, [x_2]_R, \cdots, [x_r]_R\}$ and $i \leq j \Leftrightarrow \alpha([x_i]_R) \geq \alpha([x_j]_R)$.

(4) If $\alpha([x_1]_R) \leq S(X, \underline{R}(X))$, then $X^{opt} = \underline{R}(X)$.

(5) If $\alpha([x_r]_R) \geq S(X, \overline{R}(X))$, then $X^{opt} = \overline{R}(X)$.

(6) Calculate O_i from $i = 0$ until $S(X, O_{p+1}) > \alpha([x_{p+1}]_R)$ for $p = 0, \cdots, r-1$ and set $X^{opt} = O_p$.

Theorem 2.2 and Algorithm 1 show that, by using Jaccard similarity measure, $R_{0.5}(X)$ is the optimal approximation of X in some specific cases. In addition, the computation of the optimal approximations is complex in general. In the next section, we will study optimal approximations by using a new similarity measure which is induced by symmetric difference of sets.

3. The optimal approximation of rough sets

In this section, we consider the optimal approximation of rough sets based on the following similarity measure

$$S(X,Y) = 1 - \frac{|X \triangle Y|}{|U|}, \tag{3.1}$$

where $X \triangle Y = (X - Y) \cup (Y - X)$ is the symmetric difference of X and Y. It is trivial to verify that $S(X,Y) = \frac{|X \cap Y| + |X^c \cap Y^c|}{|U|}$. That is to say, $S(X,Y)$ is the proportion of elements common to X and Y or common to X^c and Y^c in the whole universe. In addition, $S(X,Y)$ is not only related to X and Y, but also take the whole universe into consideration. Furthermore, we have the following Lemma on the relationships between S and Jaccard similarity measure S_J.

Lemma 3.1. *For any X and Y, $S(X,Y) \geq S_J(X,Y)$.*

Proof. For any X and Y, we have

$$S(X,Y) = \frac{|X \cap Y| + |X^c \cap Y^c|}{|U|} = \frac{|X \cap Y| + |X^c \cap Y^c|}{|X \cup Y| + |X^c \cap Y^c|} \geq \frac{|X \cap Y|}{|X \cup Y|} = S_J(X,Y).$$

□

Theorem 3.2. *Let (U,R) be an approximation space, $X \subseteq U$. For any definable set Y, $S(X,R_{0.5}(X)) \geq S(X,Y)$.*

Proof. Without losing generality, we assume that

$$Y \cap R_{0.5}(X) = [x_1]_R \cup [x_2]_R \cup \cdots \cup [x_k]_R,$$

$$Y - R_{0.5}(X) = [x_{i_1}]_R \cup [x_{i_2}]_R \cup \cdots \cup [x_{i_m}]_R,$$

$$R_{0.5}(X) - Y = [x_{j_1}]_R \cup [x_{j_2}]_R \cup \cdots \cup [x_{j_t}]_R,$$

$$(Y \cup R_{0.5}(X))^c = [y_1]_R \cup [y_2]_R \cup \cdots \cup [y_s]_R.$$

Then it follows that

$$\begin{aligned} & |X \cap R_{0.5}(X)| + |X^c \cap (R_{0.5}(X))^c| \\ = {} & |X \cap [x_1]_R| + \cdots + |X \cap [x_k]_R| + |X \cap [x_{j_1}]_R| + \cdots + |X \cap [x_{j_t}]_R|, \\ & +|X^c \cap [x_{i_1}]_R| + \cdots + |X^c \cap [x_{i_m}]_R| + |X^c \cap [y_1]_R| + \cdots + |X^c \cap [y_s]_R|. \\ & |X \cap Y| + |X^c \cap Y^c| \\ = {} & |X \cap [x_1]_R| + \cdots + |X \cap [x_k]_R| + |X \cap [x_{i_1}]_R| + \cdots + |X \cap [x_{i_m}]_R| \\ & +|X^c \cap [x_{j_1}]_R| + \cdots + |X^c \cap [x_{j_t}]_R| + |X^c \cap [y_1]_R| + \cdots + |X^c \cap [y_s]_R|. \end{aligned}$$

Thus we have

$$\begin{aligned} & |X \cap R_{0.5}(X)| + |X^c \cap (R_{0.5}(X))^c| - |X \cap Y| - |X^c \cap Y^c| \\ = {} & (|X \cap [x_{j_1}]_R| - |X^c \cap [x_{j_1}]_R|) + \cdots + (|X \cap [x_{j_t}]_R| - |X^c \cap [x_{j_t}]_R|) \\ & +(|X^c \cap [x_{i_1}]_R| - |X \cap [x_{i_1}]_R|) + \cdots + (|X^c \cap [x_{i_m}]_R| - |X \cap [x_{i_m}]_R|). \end{aligned}$$

For any λ, $1 \leq \lambda \leq t$, we have $\frac{|[x]_{j_\lambda} \cap X|}{|[x]_{j_\lambda}|} \geq 0.5$. So we get

$$|[x]_{j_\lambda} \cap X| \geq 0.5|[x]_{j_\lambda}| = 0.5(|[x]_{j_\lambda} \cap X| + |[x]_{j_\lambda} \cap X^c|).$$

It follows that $|[x]_{j_\lambda} \cap X| \geq |[x]_{j_\lambda} \cap X^c|$.

For any λ, $1 \leq \lambda \leq m$, we have $\frac{|[x]_{i_\lambda} \cap X|}{|[x]_{i_\lambda}|} < 0.5$. Hence we get

$$|[x]_{i_\lambda} \cap X| \leq 0.5|[x]_{i_\lambda}| = 0.5(|[x]_{i_\lambda} \cap X| + |[x]_{i_\lambda} \cap X^c|).$$

Therefore $|[x]_{i_\lambda} \cap X| \leq |[x]_{i_\lambda} \cap X^c|$. Consequently, we have $S(X,R_{0.5}(X)) \geq S(X,Y)$. □

This theorem shows that $R_{0.5}(X)$ is the optimal approximation of X with respect to the similarity measure S. The similarity measure S can be generalized to the following parameter depend form:

$$S_a(X,Y) = \frac{a(|X \cap Y| + |X^c \cap Y^c|)}{a(|X \cap Y| + |X^c \cap Y^c|) + |X \cap Y^c| + |X^c \cap Y|}, \tag{3.2}$$

where $a > 0$ is a parameter. If $a = 1$, then we have

$$S_1(X,Y) = \frac{|X \cap Y| + |X^c \cap Y^c|}{|X \cap Y| + |X^c \cap Y^c| + |X \cap Y^c| + |X^c \cap Y|} = 1 - \frac{|X \triangle Y|}{|U|} = S(X,Y).$$

The following Lemma shows the relationships between S_a and S.

Lemma 3.3. *For any X and Y, $S_a(X,Y) = \frac{aS(X,Y)}{1+(a-1)S(X,Y)}$.*

Proof. By $S(X,Y) = 1 - \frac{|X \triangle Y|}{|U|} = \frac{|U| - |X \triangle Y|}{|U|} = \frac{|X \cap Y| + |X^c \cap Y^c|}{|U|}$, we have

$$|X \cap Y| + |X^c \cap Y^c| = |U|S(X,Y).$$

Then we get

$$S_a(X,Y) = \frac{a|U|S(X,Y)}{a|U|S(X,Y) + |X \cap Y^c| + |X^c \cap Y|} = \frac{a|U|S(X,Y)}{a|U|S(X,Y) + |U| - |U|S(X,Y)}.$$

Thus it follows that $S_a(X,Y) = \frac{aS(X,Y)}{1+(a-1)S(X,Y)}$ as required. □

For the similarity measure S_a, we have:

Theorem 3.4. *Let (U,R) be an approximation space, $X \subseteq U$. For any definable set Y, $S_a(X, R_{0.5}(X)) \geq S_a(X,Y)$.*

Proof. By Theorem 3.2, we have $S(X, R_{0.5}(X)) \geq S(X,Y)$. Then we get

$$\begin{aligned} S_a(X, R_{0.5}(X)) - S_a(X,Y) &= \frac{aS(X,R_{0.5}(X))}{1+(a-1)S(X,R_{0.5}(X))} - \frac{aS(X,Y)}{1+(a-1)S(X,Y)} \\ &= \frac{a(S(X,R_{0.5}(X)) - S(X,Y))}{(1+(a-1)S(X,R_{0.5}(X)))(1+(a-1)S(X,Y))} \geq 0. \end{aligned}$$ □

This theorem shows that $R_{0.5}(X)$ is also the optimal approximation of X with respect to the parameter depend similarity measure S_a.

With the changing knowledge granularities, researchers pay their attentions to how to obtain the change laws of the uncertainty of an uncertain target set in a rough approximation space [15, 16]. Next we will analyze the change laws of the similarity between a target set X and its 0.5-approximation set in the changing knowledge spaces.

Theorem 3.5. *Let R' and R'' be two equivalence relations on U and $R'' \subseteq R'$. Then $S(X, R'_{0.5}(X)) \leq S(X, R''_{0.5}(X))$ for any $X \subseteq U$.*

Proof. Without losing generality, we assume that $U/R' = \{X_1, X_2, \cdots, X_m\}$, $U/R'' = \{Y_1, Y_2, X_2, \cdots, X_m\}$ and $X_1 = Y_1 \cup Y_2$.

(1) If $X_1 \subseteq R'_{0.5}(X)$, then $\frac{|X_1 \cap X|}{|X_1|} \geq 0.5$.

a) If $Y_1 \subseteq R''_{0.5}(X)$ and $Y_2 \subseteq R''_{0.5}(X)$, then $R'_{0.5}(X) = R''_{0.5}(X)$ and $S(X, R'_{0.5}(X)) \leq S(X, R''_{0.5}(X))$.

b) If $Y_1 \subseteq R''_{0.5}(X)$ and $Y_2 \nsubseteq R''_{0.5}(X)$, then $\frac{|Y_1 \cap X|}{|Y_1|} \geq 0.5$ and $\frac{|Y_2 \cap X|}{|Y_2|} < 0.5$.

It follows that $R''_{0.5}(X) = R'_{0.5}(X) - Y_2 = R'_{0.5}(X) \cap Y_2^c$, $R'_{0.5}(X) = R''_{0.5}(X) \cup Y_2$ and consequently

$$\begin{aligned} &|R''_{0.5}(X) \cap X| + |(R''_{0.5}(X))^c \cap X^c| - |R'_{0.5}(X) \cap X| - |(R'_{0.5}(X))^c \cap X^c| \\ &= |R''_{0.5}(X) \cap X| + |(R'_{0.5}(X) \cap Y_2)^c \cap X^c| - |(R''_{0.5}(X) \cup Y_2) \cap X| - |(R'_{0.5}(X))^c \cap X^c| \\ &= |R''_{0.5}(X) \cap X| + |(R'_{0.5}(X))^c \cap X^c| + |Y_2 \cap X^c| - |R''_{0.5}(X) \cap X| - |Y_2 \cap X| - |(R'_{0.5}(X))^c \cap X^c| \\ &= |Y_2 \cap X^c| - |Y_2 \cap X|. \end{aligned}$$

By $\frac{|Y_2 \cap X|}{|Y_2|} < 0.5$, it follows that $|Y_2 \cap X| < 0.5|Y_2| = 0.5(|Y_2 \cap X| + |Y_2 \cap X^c|)$. Then $|Y_2 \cap X| < |Y_2 \cap X^c|)$ and $|R''_{0.5}(X) \cap X| + |(R''_{0.5}(X))^c \cap X^c| - |R'_{0.5}(X) \cap X| - |(R'_{0.5}(X))^c \cap X^c| > 0$. It follows that $S(X, R'_{0.5}(X)) \leq S(X, R''_{0.5}(X))$.

c) If $Y_1 \nsubseteq R''_{0.5}(X)$ and $Y_2 \subseteq R''_{0.5}(X)$, then $S(X, R'_{0.5}(X)) \leq S(X, R''_{0.5}(X))$ can be proved similarly.

(2) If $X_1 \nsubseteq R'_{0.5}(X)$, then $\frac{|X_1 \cap X|}{|X_1|} = \frac{|Y_1 \cap X| + |Y_2 \cap X|}{|X_1|} < 0.5$.

a) If $Y_1 \nsubseteq R''_{0.5}(X)$ and $Y_2 \nsubseteq R''_{0.5}(X)$, then $R'_{0.5}(X) = R''_{0.5}(X)$ and $S(X, R'_{0.5}(X)) \leq S(X, R''_{0.5}(X))$.

b) If $Y_1 \subseteq R''_{0.5}(X)$ and $Y_2 \nsubseteq R''_{0.5}(X)$, then $\frac{|Y_1 \cap X|}{|Y_1|} \geq 0.5$ and $\frac{|Y_2 \cap X|}{|Y_2|} < 0.5$. It follows that $R''_{0.5}(X) = R'_{0.5}(X) \cup Y_1$, $R'_{0.5}(X) = R''_{0.5}(X) - Y_1 = R''_{0.5}(X) \cap Y_1^c$. Then

$$|R''_{0.5}(X) \cap X| + |(R''_{0.5}(X))^c \cap X^c| - |R'_{0.5}(X) \cap X| - |(R'_{0.5}(X))^c \cap X^c|$$
$$= |(R'_{0.5}(X) \cup Y_1) \cap X| + |(R''_{0.5}(X))^c \cap X^c| - |R'_{0.5}(X) \cap X| - |(R''_{0.5}(X) \cap Y_1^c)^c \cap X^c|$$
$$= |R'_{0.5}(X) \cap X| + |Y_1) \cap X| + |(R''_{0.5}(X))^c \cap X^c| - |R'_{0.5}(X) \cap X| - |(R''_{0.5}(X))^c \cap X^c| - |Y_1 \cap X^c|$$
$$= |Y_1) \cap X| - |Y_1 \cap X^c|$$

By $Y_1 \subseteq R''_{0.5}(X)$, we have $\frac{|Y_1 \cap X|}{|Y_1|} \geq 0.5$, it follows that $|Y_1 \cap X| \geq 0.5|Y_1| = 0.5(|Y_1 \cap X| + |Y_1 \cap X^c|)$. Thus $|Y_1 \cap X| \geq |Y_1 \cap X^c|)$ and $|R''_{0.5}(X) \cap X| + |(R''_{0.5}(X))^c \cap X^c| - |R'_{0.5}(X) \cap X| - |(R'_{0.5}(X))^c \cap X^c| > 0$. It follows that $S(X, R'_{0.5}(X)) \leq S(X, R''_{0.5}(X))$.

c) If $Y_1 \nsubseteq R''_{0.5}(X)$ and $Y_2 \subseteq R''_{0.5}(X)$, then $S(X, R'_{0.5}(X)) \leq S(X, R''_{0.5}(X))$ can be proved similarly. □

4. The attribute reduction of decision systems based on optimal approximation of rough sets

Attribute reduction is one of the core issues of the rough set theory which has been widely studied [17, 18, 19]. By deleting irrelevant or unimportant attributes under some specific conditions, approximate decision rules can be acquired from decision systems. In this section, we present an attribute reduction approach which preserves the optimal approximations of decision classes unchanged.

Definition 4.1. A *decision making information system* S can be described as $S = (U, C \cup D, V, f)$, where U is the set of objects, C is the set of conditional attributes, D is the decision attribute, V is the sets of attribute values, and $f : U \times (C \cup D) \to V$ is an information function, it specifies the attribute value of every object x in U.

Given a decision making information system $S = (U, C \cup D, V, f)$. For any $A \subseteq C$, an indiscernibility relation R^A is defined as follows:

$$R^A = \{(x, y) \in U \times U | \forall a \in A(f(x, a) = f(y, a)\} \tag{4.1}$$

Definition 4.2. Let $S = (U, C \cup D, V, f)$ be a decision system, $U/D = \{D_1, D_2, \cdots, D_n\}$, $A \subseteq C$. If $R^A_{0.5}(D_i) = R^C_{0.5}(D_i)$ for each $D_i \in U/D$, we say that A *is a optimal approximation consistent set* of S. If A is a optimal approximation consistent set and no proper subset of A is optimal approximation consistent, then A is referred to as an *optimal approximation reduction* of S.

For any $A \subseteq C$ and $x \in U$, let $M_A(x) = \{D_j | x \in R^A_{0.5}(D_j)$.

Theorem 4.3. *Let $S = (U, C \cup D, V, f)$ be a decision system, $U/D = \{D_1, D_2, \cdots, D_n\}$, $A \subseteq C$. Then*

(1) *A is a optimal approximation consistent set if and only if $M_A(x) = M_C(x)$ for each $x \in U$,*

(2) *A is a optimal approximation consistent set if and only if for any $x, y \in U$, if $M_C(x) \neq M_C(y)$, then $[x]_A \neq [y]_A$.*

Proof. (1) By $x \in R^A_{0.5}(D_j)$ if and only if $D_j \in M_A(x)$, and $x \in R^C_{0.5}(D_j)$ if and only if $D_j \in M_C(x)$, we can directly conclude (1).

(2) $\Rightarrow$: Assume that A is a optimal approximation consistent set and $x, y \in U$ with $M_C(x) \neq M_C(y)$. By (1), we have $M_A(x) = M_C(x)$ and $M_A(y) = M_C(y)$. Then it follows that $M_A(x) \neq M_A(y)$. Thus $[x]_A \neq [y]_A$.

$\Leftarrow$: For each $x \in U$, we prove $M_A(x) = M_C(x)$. Let $T = \{[y]_C | [y]_C \subseteq [x]_A\}$. Then T forms a partition of $[x]_A$.

For any $D_j \in M_C(x)$, we have $x \in R^C_{0.5}(D_j)$ and $\frac{|[x]_C \cap D_j|}{|[x]_C|} \geq 0.5$. For any $[y]_C \in T$, we have $[y]_C \subseteq [x]_A$. Then $[y]_A = [x]_A$. It follows that $M_C(x) = M_C(y)$. Thus $D_j \in M_C(y)$. It follows that $\frac{|[y]_C \cap D_j|}{|[y]_C|} \geq 0.5$ and $|[y]_C \cap D_j| \geq 0.5|[y]_C|$. Consequently,

$$\frac{|[x]_A \cap D_j|}{|[x]_A|} = \frac{\sum_{[y]_C \in T} |[y]_C \cap D_j|}{|[x]_A|} \geq \frac{\sum_{[y]_C \in T} 0.5|[y]_C|}{|[x]_A|} = 0.5\frac{\sum_{[y]_C \in T} |[y]_C|}{|[x]_A|} = 0.5.$$

It follows that $D_j \in M_A(x)$.

For any $D_j \notin M_C(x)$, we have $x \notin R^C_{0.5}(D_j)$. For any $[y]_C \in T$, we have $[y]_C \subseteq [x]_A$. Then $[y]_A = [x]_A$. It follows that $M_C(x) = M_C(y)$. Thus $D_j \notin M_C(y)$. It follows that $\frac{|[y]_C \cap D_j|}{|[y]_C|} < 0.5$ and $|[y]_C \cap D_j| < 0.5|[y]_C|$. Consequently,

$$\frac{|[x]_A \cap D_j|}{|[x]_A|} = \frac{\sum_{[y]_C \in T} |[y]_C \cap D_j|}{|[x]_A|} < \frac{\sum_{[y]_C \in T} 0.5|[y]_C|}{|[x]_A|} = 0.5\frac{\sum_{[y]_C \in T} |[y]_C|}{|[x]_A|} = 0.5.$$

It follows that $D_j \notin M_A(x)$. So $M_A(x) = M_C(x)$ as required. □

Given a decision system $S = (U, C \cup D, V, f)$. For any $x, y \in U$, let $E_{xy} = \{a \in C | f(x, a) \neq f(y, a)\}$ if $M_C(x) \neq M_C(y)$, and $E_{xy} = C$ otherwise.

Theorem 4.4. *Let $S = (U, C \cup D, V, f)$ be a decision system, $A \subseteq C$. A is a optimal approximation consistent set if and only if $A \cap E_{xy} \neq \varnothing$ for any $x, y \in U$ with $M_C(x) \neq M_C(y)$.*

Proof. Assume that A is a optimal approximation consistent set. If $x, y \in U$ with $M_C(x) \neq M_C(y)$, then by Theorem 4.3, we have $[x]_A \neq [y]_A$. It follows that there exists $a \in A$ such that $f(x, a) \neq f(y, a)$.Thus $A \cap E_{xy} \neq \varnothing$.

Conversely, assume that $x, y \in U$ and $M_C(x) \neq M_C(y)$. It follows that $A \cap E_{xy} \neq \varnothing$. Then there exists $a \in A$ such that $a \in E_{xy}$. By $f(x, a) \neq f(y, a)$, we conclude that $[x]_A \neq [y]_A$. By Theorem 4.3, A is a optimal approximation consistent set. □

The logical formula $\wedge_{M_C(x) \neq M_C(y)} (\vee_{a \in E_{xy}} a)$ is called an *optimal approximation discernibility function* of S. By using the technique of discernibility matrix and discernibility function presented in [19], we have the following method to compute attribute reductions.

Theorem 4.5. *Let $S = (U, C \cup D, V, f)$ be a decision system. The minimal disjunctive normal form of optimal approximation discernibility function is*

$$\wedge_{M_C(x) \neq M_C(y)} (\vee_{a \in E_{xy}} a) = \vee_{1 \leq k \leq t} (\wedge_{1 \leq s \leq q_k} a_{ks}) \tag{4.2}$$

Let $A_k = \{a_{k1}, a_{k2}, \cdots, a_{kq_k}\}$. Then $\{A_1, A_2, \cdots, A_t\}$ is just the set of all optimal approximation reductions of S.

Example 4.6. Suppose that a decision system $S = (U, C \cup D, V, f)$ is given in Table 1, $U = \{x_i | 1 \le i \le 8\}$, $C = \{a_1, a_2, a_3, a_4\}$ and D is the decision attribute. We

TABLE 1. A decision formal context

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
a_1	0	0	1	2	0	0	0	2
a_2	0	0	0	1	1	2	1	1
a_3	0	0	0	0	0	1	1	0
a_4	0	1	0	0	0	0	1	1
D	0	0	1	1	0	1	1	0

denote $U/D = \{D_1, D_2\}$, where $D_1 = \{x_1, x_2, x_5, x_8\}$, $D_2 = \{x_3, x_4, x_6, x_7\}$. Let $A = \{a_3, a_4\}$. It follows that $U/R^A = \{\{x_1, x_3, x_4, x_5\}, \{x_2, x_8\}, \{x_6\}, \{x_7\}\}$. It follows that $R^A_{0.5}(D_1) = \{x_1, x_2, x_3, x_4, x_5, x_8\}$ and $R^A_{0.5}(D_2) = \{x_1, x_3, x_4, x_5, x_6, x_7\}$ are optimal approximations of D_1 and D_2 respectively.

To compute attribute reductions, we have $M_C(x_1) = M_C(x_2) = M_C(x_5) = M_C(x_8) = \{D_1\}$ and $M_C(x_3) = M_C(x_4) = M_C(x_6) = M_C(x_7) = \{D_2\}$. By Theorem 4.4, we compute the discernibility attribute sets and present them in Table 2.

TABLE 2. The discernibility matrix

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1			a_1	a_1a_2		a_2a_3	$a_2a_3a_4$	
x_2			a_1a_4	$a_1a_2a_4$		$a_2a_3a_4$	a_2a_3	
x_3					a_1a_2			$a_1a_2a_4$
x_4					a_1			a_4
x_5						a_2a_3	a_3a_4	
x_6								$a_1a_2a_3a_4$
x_7								a_1a_3

The discernibility function of S is:
$a_1 \wedge (a_2 \vee a_3) \wedge a_4 = (a_1 \wedge a_2 \wedge a_4) \vee (a_1 \wedge a_3 \wedge a_4)$.
Thus there are two attribute reductions: $\{a_1, a_2, a_4\}$ and $\{a_1, a_3, a_4\}$.

5. CONCLUDING REMARKS

In the application of rough sets in knowledge discovery in information systems, the optimal approximations may provide more accurate decision rules. In this paper, we present the optimal approximations of rough sets in an approximation space by using similarity measures induced by symmetric difference. The variation rules of similarity degrees between the target concepts and their optimal approximations in different granularity spaces are surveyed. Furthermore, the attribute reduction approaches for decision systems which preserve the optimal approximations of decision classes unchanged are presented. In further study, the optimal approximations and their applications in some other approximation spaces, for example, covering approximation space and fuzzy rough approximation space, will be investigated.

References

[1] Z. Pawlak, Rough Sets, International Journal of Computer and Information Science 11 (1982) 341–356.
[2] S. Ayub, W. Mahmood, M. Shabir and R. Gu, A study of modules of fractions under fuzzy sets and soft sets, New Mathematics and Natural Computation 120 (1) (2024) 247–271.
[3] R. Gul, S. Ayub, M. Shabir, A. Y. Hummdi, A. Aljaedi and Z. Bassfar, A novel approach for fuzzification of rough sets based on fuzzy preference relation: properties and application to medicine selection problem, IEEE Access (2024) 67681–67703.
[4] A. Jankowski, A. Skowron and R. W. Swiniarski, Perspectives on uncertainty and risk in rough sets and interactive rough-granular computing, Fundamenta Informatica 129 (1-2) (2014) 69–84.
[5] R. S. Kanwal, S. M. Qurashi, R. Gul, A. M. A. Ellatif, T. M. Alsham and F. Tufail, New insights into rough approximations of a fuzzy set inspired by soft relations with decision making applications, AIMS Mathematics 10 (4) (2025) 9637–9673.
[6] Z. Pawlak, Rough sets and intelligent data analysis, Inform. Sci. 147 (2002) 1–12.
[7] Z. Bonikowaski, Algebric structures of rough sets, Fuzzy Sets and Knowledge Discovery (1995) 242–247.
[8] D. Dubois and H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems 17 (2-3) (1990) 191–209.
[9] W. Ziarko, Variable precision rough set model, Journal of Computer and System Science 46 (1993) 39–59.
[10] Y. Y. Yao, Probabilistic rough set approximations, International Journal of Approximate Reasoning 49 (2) (2008) 255–271.
[11] R. Janicki and A. Lenarcic, Optimal approximations with rough sets, in: P. Lingras, M. Wolski, C. Cornelis, S. Mitra, P. Wasilewski(eds), Proc. of Rough Sets and Knowledge Technology, RSKT 2013, Lecture Notes in Artificial Intelligence 8171 (2013) 87–98.
[12] R. Janicki and A. Lenarcic, Optimal approximations with rough sets and similarities in measure spaces, International Journal of Approximate Reasoning, 71 (2016) 1–14.
[13] Q. H. Zhang, G. Y. Wang and Y. Xiao, Approximation Sets of Rough Sets, Journal of Software 23 (2012) 1745–1759.
[14] Q. H. Zhang, J. Wang, G. Y. Wang and H. Yu, The approximation set of a vague set in rough approximation space, Inform. Sci. 300 (2015) 1–19.
[15] Q. H. Zhang, Y. L. Guo and Y. Xiao, Attribute reduction based on approximation set of rough set, Journal of Computational Information Systems 10 (16) (2014) 6859–6866.
[16] Q. H. Zhang, P. Zhang and G. Y. Wang, Research on approximation set of rough set based on fuzzy similarity, Journal of Intelligent and Fuzzy Systems 32 (2017) 2549–2562.
[17] M. Kryszkiewicz, Comparative study of alternative of type of knowledge reduction in inconsistent systems, International Jounal of Intelligent Systems 16 (2001) 105–120.
[18] J. S. Mi, W. Z. Wu and W. X. Zhang, Approaches to knowledge reduction based on variable precision rough set model, Information Sciences 159 (2004) 255–272.
[19] A. Skowron and C. Rauszer, The discernibility matrices and functions in information systems, in: R. Slowinski(Ed.), Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory, Kluwer, Academic, Dordrecht, 1992, pp 331–362.

Rui Qin (zarathustra8711@gmail.com)
Corporate Development and Strategic MA, Shanghai Microport CardioFlow Medtech Co., Ltd., 201203, Shanghai, China

Zhi ZhengXiao (constantinozheng@gmail.com)
Exotic Equity Derivative Sales Trading, Banco de Bilbao Vizcaya Arengentaria, Madrid 28050, Spain

Keyun Qin (keyunqin@263.net)
College of mathematics, Southwest Jiaotong University, 610031, Chengdu, China