Computing and Informatics, Vol. 33, 2014, 633–651

KNOWLEDGE GRANULATION, ROUGH ENTROPY AND UNCERTAINTY MEASURE IN INCOMPLETE FUZZY INFORMATION SYSTEM

Qiang Wu

Department of Computer Science and Technology
Shaoxing University
Shaoxing, 312000
Zhejiang, People's Republic of China
e-mail: cswq@usx.edu.cn

Abstract. Many real world problems deal with ordering of objects instead of classifying objects, although most of research in data analysis has been focused on the latter. One of the extensions of classical rough sets to take into account the ordering properties is dominance-based rough sets approach which is mainly based on substitution of the indiscernibility relation by a dominance relation. In this paper, we address knowledge measures and reduction in incomplete fuzzy information system using the approach. Firstly, new definitions of knowledge granulation and rough entropy are given, and some important properties of them are investigated. Then, dominance matrix about the measures knowledge granulation and rough entropy is obtained, which could be used to eliminate the redundant attributes in incomplete fuzzy information system. Lastly, a matrix algorithm for knowledge reduction is proposed. An example illustrates the validity of this method and shows the method is applicable to complex fuzzy system. Experiments are also made to show the performance of the newly proposed algorithm.

Keywords: Incomplete fuzzy information system, dominance relation, knowledge granulation, rough entropy, knowledge reduction

1 INTRODUCTION

The rough sets theory proposed by Pawlak [10, 11, 12] has been studied intensively and obtained results have played an important role in data analysis and knowledge processing. Because of its usefulness, rough set approaches have been used in pattern recognition [14], medical informatics [6], market decision analysis [15], kansei engineering and so on [25]. The classical rough sets are defined using the indiscernibility relation, i.e., an equivalence relation. This implies that it can only be used to process information systems with discrete data. However, in the real world, we may face cases when some attribute values are fuzzy. For example, as with temperature, wind and outlook etc., the weather forecast attributes such as decision tables, its temperature would be able to get hot, medium, cool, and so fuzzy.

At present, many researchers extended the classical rough set model in fuzzy environment [3, 9, 19, 17]. They put forward a variety of fuzzy rough sets and rough fuzzy set model. It is noteworthy that these works just assume that all the attributes of fuzzy information systems are conventional properties, not taking into account the different attributes of the relationship between the values of the partial order. To address this problem, Greco et al. introduced the concept of dominance relation into fuzzy information system [4, 21]. It is proposed that the rough fuzzy set model is based on dominance relation [3]. This rough fuzzy approach is different from the classical rough fuzzy and fuzzy rough techniques, because it is based on different fuzzy membership of the partial order relation. Concepts of knowledge granulation, knowledge entropy and knowledge uncertainty measure are given in ordered information systems by Xu et al. [20], and some important properties of them are investigated. Yang and Wei et al. [18] present a general framework for the study of dominance based rough set model in the incomplete fuzzy information systems. Soon, their another paper is to further investigate the dominance based rough set in incomplete interval-valued information system, which contains both incomplete and imprecise evaluations of objects. However, they are confined to traditional methods for the measure of knowledge.

In this paper, we investigate incomplete fuzzy information system (IFIS) with some unknown attribute values. We believe that the unknown attribute value is just missing, but it is there. Each individual object has the potential to yield complete information by appropriate means. Therefore, in IFIS, the unknown attribute values can be considered as any other known attribute values are comparable. According to this explanation, a new fuzzy rough set model is built. Based on this model, we address knowledge granulation and rough entropy of rough set in IFIS, thus to knowledge reduction.

In the next section, IFIS and its extended dominance relations are reviewed. In Section 3, knowledge granulation is defined and rough entropy of rough set based on dominance relations is proposed in IFIS, and we get a number of conclusions. In Section 4, after giving knowledge reduction algorithm based on dominance matrix, an example shows all reducts are enumerated by the matrices associated with IFIS. In Section 5, our method is compared with the existing algorithms by the UCI database for attribute reduction. Finally, we present our conclusions in Section 6.

2 BASIC NOTIONS

In this section, we introduce some basic terminology and notations which will be used throughout the paper.

2.1 Incomplete Fuzzy Information System

The fuzzy set theory can be thought of as an extension of traditional crisp sets in which each element must either be or not be in a set. Formally, the process by which individuals from a universal set $\mho$ are determined to be either members or non-members of a crisp set can be defined by a characteristic or discrimination function. This kind of function can be generalized such that the values assigned to the elements of the universal set fall within specified ranges, referred to as the membership function $\mu_A(v)$, by which a fuzzy set A is usually defined. This function is represented by

$$\mu_A : \mho \to [0, 1]$$

where $[0, 1]$ denotes the interval of real numbers from 0 to 1, inclusive. The function can also be generalized to any real interval and is not restricted to $[0, 1]$.

The scalar cardinality of a fuzzy set A defined on a finite universal set $\mho$ is the summation of the membership grades of all the elements of $\mho$ in A. Thus,

$$|A| = \sum_{v \in \mho} \mu_A(v).$$

A complete information system S is a 4-tuple $S = \langle U, AT, V, f \rangle$, where U is a non-empty finite set of objects called the universe and AT is a non-empty finite set of attributes such that $a : U \to V_a$ for any $a \in AT$, i.e. $f(x, a) \in V_a$, where V_a is called the domain of attribute a. When the precise values of some of the attributes in an information system are not known, i.e. are missing or known partially, such a system is called an incomplete information system.

In an incomplete information system, the values of $a \in AT$ are subsets of V_a; if $f(x, a)$ is a normalized fuzzy subset of V_a for each $x \in U$, then the system is called an incomplete fuzzy information system (IFIS) and is still denoted without confusion by $S = \langle U, AT, V, f \rangle$. The fuzzy set $f(x, a) \in F(V_a)$ may be seen as a possibility distribution on V_a, and the quantity $f(x, a)_{(v)}, v \in V_a$ thus represents the degree of possibility of state v. We assume here that at least one of the states of V_a is fully possible for each $a \in AT$, i.e. $v \in V_a$ such that $f(x, a)_{(v)} = 1$. Thus some states are more possible than others.

The IFISs are generalizations of incomplete information systems and complete information systems. A complete information system is an extreme form of an IFIS with partial knowledge in which the knowledge of each object with respect to each attribute is complete, i.e. for some $v \in V_a, f(x, a)_{(v)} = 1$ and $f(x, a)_{(s)} = 0, \forall s \neq v$, whereas an incomplete information system with missing values is another extreme form of IFIS with partial knowledge in which the knowledge of some objects (with missing values) is complete ignorance, i.e. $f(x,a)_{(v)} = 1, \forall v \in V_a$.

Example 1. [19] Table 1 illustrates an exemplary IFIS $S = \langle U, AT, V, f\rangle$, where $U = \{1, 2, \ldots, 8\}$, $AT = \{a, b, c\}, V_a = \{H, N, L\}, V_b = \{R, S, T\}, V_c = \{m, n, p\}$.

According to the above statement, it is easy to see that attribute values $a(7), b(3)$ and $c(3)$ are missing, $a(4)$ does not take H as value, $b(6)$ does not take T as value, and $c(6)$ does not take n as value. The knowledge of $a(1), a(3), a(5), b(1), b(4), b(7)$ and $c(4)$ is complete.

U	a	b	c
1	$1/H$	$1/S$	$1/m + 0.1/p$
2	$1/H + 0.3/N$	$1/S + 0.4/T$	$1/m + 0.5/n$
3	$1/H$	$1/R + 1/S + 1/T$	$1/m + 1/n + 1/p$
4	$1/N + 1/L$	$1/T$	$1/n$
5	$1/N$	$1/R + 0.6/S$	$0.7/m + 1/p$
6	$0.1/H + 0.4/N + 1/L$	$1/R + 1/S$	$1/m + 1/p$
7	$1/H + 1/N + 1/L$	$1/T$	$1/n + 0.2/p$
8	$0.8/N + 1/L$	$0.7/S + 1/T$	$0.8/m + 1/n$

Table 1. Exemplary IFIS

In IFIS, the membership grades of two fuzzy sets can be used to compare the objects. Therefore, we can define the scalar cardinality of attributes to the system:

$$|f(x,a)| = \sum_{v \in V_a} f(x,a)_{(v)}, f(x,a)_{(v)} \neq 1 \quad \text{and} \quad f(x,a)_{(v)} \neq 0.$$

When the precise values for some of the objects on some fuzzy attributes are not known, i.e. unknown values (symbol "$*$" is used to express unknown value), the fuzzy information system is still referred to as an IFIS. In this paper, an IFIS is still recorded as $S = \langle U, AT, V, f\rangle$, at this time $V = [0,1] \cup \{*\}$.

Table 2 is an IFIS, of which $U = \{x_1, x_2, \ldots, x_{10}\}$, attribute set $AT = \{a_1, a_2, a_3, a_4\}$.

2.2 Dominance Relation

Definition 1. Let $S = \langle U, AT, V, f\rangle$ be an IFIS, $B \subseteq AT$. The dominance relation in terms of B is defined as $R_B^{\gtrless} = \{(x,y) \in U^2 | \forall a \in B, f(x,a) \geq f(y,a) \vee f(x,a) = * \vee f(y,a) = *\}$. $[x_i]_B^{\gtrless}$ is called a dominance class of object x_i, if $[x_i]_B^{\gtrless} = \{x_j \in U | (x_j, x_i) \in R_B^{\gtrless}\} = \{x_j \in U | \forall a \in B, f(x_j,a) \geq f(x_j,a) \vee f(x_i,a) = *\}$, $U/R_B^{\gtrless} = \{[x_i]_B^{\gtrless} | x_j \in U\}$. $U/R_B^{\gtrless}$ is a classification for the object set on the attribute set B.

It can be verified easily that the dominance relation satisfies the properties as follows.

U	a_1	a_2	a_3	a_4
x_1	0.9	$*$	0.2	0.7
x_2	0.9	0.2	0.2	0.1
x_3	0.1	0.1	0.1	0.9
x_4	0.0	0.9	$*$	0.8
x_5	0.1	0.1	1.0	0.8
x_6	$*$	0.2	0.9	0.1
x_7	0.0	0.1	0.9	0.2
x_8	0.9	0.9	0.1	1.0
x_9	0.8	0.4	1.0	1.0
x_{10}	0.0	1.0	0.2	$*$

Table 2. An IFIS with unknown values

Property 1. Let S be an IFIS and $B, C \subseteq AT$. Then

1. $R_B^{\gtrsim}$ is reflexive and transitive, not necessarily symmetric. So, it is no longer an equivalence relation;
2. $B \subseteq C \Rightarrow R_C^{\gtrsim} \subseteq R_B^{\gtrsim}$;
3. $B \subseteq C \Rightarrow [x_i]_C^{\gtrsim} \subseteq [x_i]_B^{\gtrsim}$;
4. $x_j \in [x_i]_C^{\gtrsim} \Rightarrow [x_j]_B^{\gtrsim} \subseteq [x_i]_C^{\gtrsim}$;
5. $[x_j]_B^{\gtrsim} \subseteq [x_i]_B^{\gtrsim} \Leftrightarrow f(x_j, a) = f(x_i, a)(\forall a \in B)$;
6. $|[x_i]_B^{\gtrsim}| \geq 1$, for any $x_i \in U$, where $|\cdot|$ denotes cardinality of $[x_i]_B^{\gtrsim}$;
7. $U/R_B^{\gtrsim}$ constructs a cover of U, that is, $[x_i]_B^{\gtrsim} \neq \emptyset$ and $\cup_{x_i \in U}[x_i]_B^{\gtrsim} = U$ for any $x_i \in U$.

In order to better describe the classification in IFIS, the following definition is given.

Definition 2. Let $S = \langle U, AT, V, f \rangle$ be an IFIS, and $B, C \subseteq AT$.

1. $U/R_B^{\gtrsim} = U/R_C^{\gtrsim}$, if $[x_i]_B^{\gtrsim} \subseteq [x_i]_C^{\gtrsim}$ and $[x_i]_B^{\gtrsim} \supseteq [x_i]_C^{\gtrsim}$ for any $x_i \in U$;
2. $U/R_B^{\gtrsim} \subseteq U/R_C^{\gtrsim}$, if $[x_i]_B^{\gtrsim} \subseteq [x_i]_C^{\gtrsim}$ for any $x_i \in U$. It is said $U/R_B^{\gtrsim}$ is finer than $U/R_C^{\gtrsim}$;
3. $U/R_B^{\gtrsim} \subset U/R_C^{\gtrsim}$, if $[x_i]_B^{\gtrsim} \subseteq [x_i]_C^{\gtrsim}$ for any $x_i \in U$ and $[y]_B^{\gtrsim} \neq [y]_C^{\gtrsim}$. It is said $U/R_B^{\gtrsim}$ is indeed finer than $U/R_C^{\gtrsim}$.

Clearly, by Property 1 and the above definition, you can get $U/R_{AT}^{\gtrsim} \subseteq U/R_B^{\gtrsim}$ immediately for $S = \langle U, AT, V, f \rangle$ and $B \subseteq AT$.

In terms of dominance classes of B, the pair of lower and upper approximation operators can be defined by

$$\begin{aligned} \underline{R_B^{\gtrsim}}(X) &= \{x_i \in U | [x_i]_B^{\gtrsim} \subseteq X\} \\ \overline{R_B^{\gtrsim}}(X) &= \{x_i \in U | [x_i]_B^{\gtrsim} \cap X \neq \emptyset\}. \end{aligned}$$

An element $x \in U$ belongs to the lower approximation of X if all its dominance elements belong to X. It belongs to the upper approximation of X if at least one of its dominance elements belongs to X.

Like Pawlak approximation space, it is also to have many similar properties. For details, please refer to literature [3].

Example 2. Table 2 gives an IFIS. As a result, by the definition of the dominance relation, we have

$$\begin{aligned}
[x_1]^{\geq}_{AT} &= \{x_1\},\\
[x_2]^{\geq}_{AT} &= \{x_1, x_2, x_6\},\\
[x_3]^{\geq}_{AT} &= \{x_3, x_8, x_9\},\\
[x_4]^{\geq}_{AT} &= \{x_4, x_8, x_9\},\\
[x_5]^{\geq}_{AT} &= \{x_5, x_9\},\\
[x_6]^{\geq}_{AT} &= \{x_4, x_6, x_9\},\\
[x_7]^{\geq}_{AT} &= \{x_4, x_5, x_7, x_9\},\\
[x_8]^{\geq}_{AT} &= \{x_8\},\\
[x_9]^{\geq}_{AT} &= \{x_9\},\\
[x_{10}]^{\geq}_{AT} &= \{x_1, x_{10}\}
\end{aligned}$$

Furthermore, if $A = \{a_1, a_2, a_3\}, B = \{a_1, a_2\}$ then

$$\begin{aligned}
[x_1]^{\geq}_{A} &= [x_2]^{\geq}_{A} = \{x_1, x_2, x_6\},\\
[x_3]^{\geq}_{A} &= \{x_1, x_2, x_3, x_5, x_6, x_8, x_9\},\\
[x_4]^{\geq}_{A} &= \{x_1, x_4, x_8, x_9\},\\
[x_5]^{\geq}_{A} &= \{x_5, x_9\},\\
[x_6]^{\geq}_{A} &= \{x_4, x_6, x_9\},\\
[x_7]^{\geq}_{A} &= \{x_4, x_5, x_7, x_9\},\\
[x_8]^{\geq}_{A} &= \{x_1, x_8\},\\
[x_9]^{\geq}_{A} &= \{x_9\},\\
[x_{10}]^{\geq}_{A} &= \{x_1, x_{10}\},
\end{aligned}$$

as well as

$$\begin{aligned}
[x_1]^{\geq}_{B} &= [x_2]^{\geq}_{B} = \{x_1, x_2, x_6, x_8\},\\
[x_3]^{\geq}_{B} &= \{x_1, x_2, x_3, x_5, x_6, x_8, x_9\},\\
[x_4]^{\geq}_{B} &= \{x_1, x_4, x_8, x_9\},
\end{aligned}$$

$$\begin{aligned}
[x_5]_B^{\geqslant} &= \{x_1, x_2, x_3, x_5, x_6, x_8, x_9\}, \\
[x_6]_B^{\geqslant} &= \{x_1, x_2, x_4, x_6, x_8, x_9, x_{10}\}, \\
[x_7]_B^{\geqslant} &= \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}, \\
[x_8]_B^{\geqslant} &= \{x_1, x_8\}, \\
[x_9]_B^{\geqslant} &= \{x_1, x_8, x_9\}, \\
[x_{10}]_B^{\geqslant} &= \{x_1, x_{10}\}.
\end{aligned}$$

Obviously, we can see from the above that $U/R_{AT}^{\geqslant} \subseteq U/R_A^{\geqslant} \subseteq U/R_B^{\geqslant}$, that is, $U/R_A^{\geqslant}$ to be finer than $U/R_B^{\geqslant}$, $U/R_{AT}^{\geqslant}$ should be finer than $U/R_A^{\geqslant}$.

3 KNOWLEDGE GRANULATION AND ROUGH ENTROPY

3.1 Knowledge Granulation

In this section, knowledge granulation and rough entropy in IFIS are introduced. They have some very useful properties.

Definition 3. Let $S = \langle U, AT, V, f \rangle$ be an incomplete fuzzy information system. The granulation of knowledge $B \subseteq AT$ is defined as follows:

$$GK(B) = \sum_{i=1}^{n} \frac{|[x_i]_B^{\geqslant}|}{|U|^2}.$$

Example 3. For Example 2, one can calculate the granulation of knowledge A and B: $GK(A) = 1/100(3 + 3 + 7 + 4 + 2 + 3 + 4 + 2 + 1 + 2) = 0.31, GK(B) = 1/100(4 + 4 + 7 + 4 + 7 + 7 + 10 + 2 + 3 + 2) = 0.50$.

A simple result can be easily obtained by Definition 3.

Property 2. Let $S = \langle U, AT, V, f \rangle$ be an IFIS, $B \subseteq AT$. Then,

1. $GK(B)$ achieves its maximum value 1 if and only if $U/R_B^{\geqslant} = U$;
2. $GK(B)$ achieves its minimum value $1/U$ if and only if $U/R_B^{\geqslant} = \{x_1, x_2, \ldots, x_{|U|}\}$.

This shows that we have $1/U \leq GK(B) \leq 1$ when $R_B^{\geqslant}$ is a dominance relation with respect to B. The granulation can also represent the discernibility ability of knowledge: the smaller $GK(B)$, the stronger its discernibility ability.

Theorem 1. Let $S = \langle U, AT, V, f \rangle$ be an IFIS, and $B_1, B_2 \subseteq AT$. If $U/R_{B_1}^{\geqslant} \subset U/R_{B_2}^{\geqslant}$, then $GK(B_1) < GK(B_2)$.

Proof. Since $U/R_{B_1}^{\geqslant} \subset U/R_{B_2}^{\geqslant}$, we have that $[x_i]_{B_1}^{\geqslant} \subseteq [x_i]_{B_2}^{\geqslant}$ for any $x_i \in U$ and there exists some $x_j \in U$ such that $|[x_j]_{B_1}^{\geqslant}| < |[x_j]_{B_2}^{\geqslant}|$. According to Definition 3 and Property 1, it holds that $\sum_{i=1}^{n} \frac{|[x_i]_{B_1}^{\geqslant}|}{|U|^2} < \sum_{i=1}^{n} \frac{|[x_i]_{B_2}^{\geqslant}|}{|U|^2}$, that is, $GK(B_1) < GK(B_2)$. This completes the proof. □

Corollary 1. Let $S = \langle U, AT, V, f\rangle$ be an IFIS, and $B_1, B_2 \subseteq AT$. If $B_1 \subseteq B_2$, then $GK(B_1) < GK(B_2)$.

Corollary 1 states that the granulation decreases as dominance classes become smaller through finer classification.

Theorem 2. Let $S = \langle U, AT, V, f\rangle$ be an IFIS, and $B_1, B_2 \subseteq AT$. If $U/R^{\geq}_{B_1} = U/R^{\geq}_{B_2}$, then $GK(B_1) = GK(B_2)$.

Proof. $|[x_i]^{\geq}_{B_1}| = |[x_i]^{\geq}_{B_2}|$ for any $x_i \in B_1, B_2$, since $U/R^{\geq}_{B_1} = U/R^{\geq}_{B_2}$. By Definition 3 we can deduce that $GK(B_1) = GK(B_2)$. This completes the proof. □

Theorem 3. Let $S = \langle U, AT, V, f\rangle$ be an IFIS, and $B_1, B_2 \subseteq AT$. If $U/R^{\geq}_{B_1} \subseteq U/R^{\geq}_{B_2}$ and $GK(B_1) = GK(B_2)$, then $U/R^{\geq}_{B_1} = U/R^{\geq}_{B_2}$.

Proof. Since $GK(B_1) = GK(B_2)$, we have that $\sum_{i=1}^{n} \frac{|[x_i]^{\geq}_{B_1}|}{|U|^2} = \sum_{i=1}^{n} \frac{|[x_i]^{\geq}_{B_2}|}{|U|^2}$ for any $x_i \in U$. From $U/R^{\geq}_{B_1} \subseteq U/R^{\geq}_{B_2}$, it follows that $[x_i]^{\geq}_{B_1} \subseteq [x_i]^{\geq}_{B_2}$ and $1 \leq |[x_i]^{\geq}_{B_1}| \leq |[x_i]^{\geq}_{B_2}|$ such that $|[x_i]^{\geq}_{B_1}| = |[x_i]^{\geq}_{B_2}|$. Therefore $U/R^{\geq}_{B_1} = U/R^{\geq}_{B_2}$. This completes the proof. □

Corollary 2. Let $S = \langle U, AT, V, f\rangle$ be an IFIS, and $B_1, B_2 \subseteq AT$. If $B_1 \subseteq B_2$ and $GK(B_1) = GK(B_2)$, then $U/R^{\geq}_{B_1} = U/R^{\geq}_{B_2}$.

Example 4. We can obtain $U/R^{\geq}_{A} = U/R^{\geq}_{B}$ in Example 2 and calculate the granulation $GK(A) = 0.31, GK(B) = 0.50$, that is, $GK(A) < GK(B)$ in Example 3. However, if let $B' = \{a_1\}$ and $B'' = \{a_2\}$, then

$$
\begin{aligned}
[x_1]^{\geq}_{B'} &= [x_2]^{\geq}_{B'} = \{x_1, x_2, x_6, x_8\},\\
[x_3]^{\geq}_{B'} &= \{x_1, x_2, x_3, x_5, x_6, x_8, x_9\},\\
[x_4]^{\geq}_{B'} &= [x_6]^{\geq}_{B'} = [x_7]^{\geq}_{B'} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\},\\
[x_5]^{\geq}_{B'} &= \{x_1, x_2, x_3, x_4, x_5, x_6, x_8, x_9\},\\
[x_8]^{\geq}_{B'} &= \{x_1, x_2, x_6, x_8\},\\
[x_9]^{\geq}_{B'} &= \{x_1, x_2, x_6, x_8, x_9\}.\\
[x_1]^{\geq}_{B''} &= \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\},\\
[x_2]^{\geq}_{B''} &= [x_6]^{\geq}_{B''} = [x_8]^{\geq}_{B''} = \{x_1, x_2, x_4, x_6, x_8, x_9, x_{10}\},\\
[x_3]^{\geq}_{B''} &= [x_5]^{\geq}_{B''} = [x_7]^{\geq}_{B''} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\},\\
[x_4]^{\geq}_{B''} &= \{x_1, x_4, x_8, x_{10}\},\\
[x_{10}]^{\geq}_{B''} &= \{x_1, x_{10}\}.
\end{aligned}
$$

One can get that $GK(B') = 0.71$ and $GK(B'') = 0.69$. It is easy to see that $U/R^{\geq}_{B'} \supseteq U/R^{\geq}_{B''}$ is not true, despite $GK(B') > GK(B'')$. So, the converse of Theorem 1 is not true.

3.2 Rough Entropy

The concept of rough entropy has been introduced in rough set, rough relational databases and incomplete information system [1, 7, 8]. Now we introduce a new definition of rough entropy of knowledge in IFIS.

Definition 4. Let $S = \langle U, AT, V, f \rangle$ be an IFIS, $B \subseteq AT$. The roughness of rough set $X \subseteq U$ with respect to knowledge B is defined by

$$\beta_B = 1 - \frac{|\underline{R_B^{\geq}}(X)|}{|\overline{R_B^{\geq}}(X)|}.$$

It can be seen from Definition 4 that the roughness of rough set is between 0 and 1. We can take this fact one step farther to get the following theorem.

Theorem 4. Let $S = \langle U, AT, V, f \rangle$ be an IFIS, and $B_1, B_2 \subseteq AT$. If $U/R_{B_1}^{\geq} \subseteq U/R_{B_2}^{\geq}$, then $\beta_{B_1} \leq \beta_{B_2}$ for any $X \subseteq U$.

Example 5. Considering a rough set of $X = \{x_1, x_2, x_8, x_9\}$ about knowledge $A = \{a_1, a_2, a_3\}$ and $B = \{a_1, a_2\}$ in Example 2, respectively, we have that $\underline{R_B^{\geq}}(X) = \{x_8, x_9\}$, $\overline{R_B^{\geq}}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}, \beta_A(X) = 1 - 2/10 = 8/10$; $\underline{R_B^{\geq}}(X) = \{x_8, x_9\}, \overline{R_B^{\geq}}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}, \beta_B(X) = 1 - 2/10 = 8/10$. Thus, $\beta_A(X) = \beta_B(X)$.

Notice that, in Example 5, $B \subset A$, but the same roughness can be obtained for the rough set X. Therefore, it is necessary for us to introduce a more accurate measure for rough sets in IFIS.

Definition 5. let $S = \langle U, AT, V, f \rangle$ be an IFIS, $B \subseteq AT$. The rough entropy of $X \subseteq U$ about knowledge B is defined as follows: $E_B(X) = \beta_B(X) GK(B)$.

Example 6. In Example 5, rough entropy of $X = \{x_1, x_2, x_8, x_9\}$ about knowledge $A = \{a_1, a_2, a_3\}$ and $B = \{a_1, a_2\}$ are: $E_A(X) = \beta_A(X) GK(A) = \frac{8}{10^3} \times 31 = 0.248$, $E_B(X) = \beta_B(X) GK(B) = \frac{8}{10^3} \times 50 = 0.40$. Then, we have that $E_A(X) < E_B(X)$.

It is clarified that this rough entropy of rough set is more accurate than the others.

Theorem 5. Let $S = \langle U, AT, V, f \rangle$ be an IFIS, and $B_1, B_2 \subseteq AT$. If $U/R_{B_1}^{\geq} = U/R_{B_2}^{\geq}$, then $E_{B_1} \leq E_{B_2}$ for any $X \subseteq U$.

Proof. This follows immediately from Theorem 4 and Definition 5. This completes the proof. □

Corollary 3. Let $S = \langle U, AT, V, f \rangle$ be an IFIS, $B_1, B_2 \subseteq AT$. If $B_2 \subseteq B_1$, then $E_{B_1} \leq E_{B_2}$ for any $X \subseteq U$.

Obviously, rough entropy can represent the discernibility ability of knowledge: the smaller $E(X)$, the stronger its discernibility ability.

4 DOMINANCE MATRIX AND MATRIX ALGORITHM FOR REDUCTION OF KNOWLEDGE

Attribute reduction is a very important part of the rough set theory. However, the cost of reduct computation is highly influenced by the size of object set and attribute set, so enhancing the reduct computation efficiency is one of the major problems. This section proposes a reduction algorithm reducing the time cost by dominance matrix.

4.1 Dominance Matrix

Definition 6. Let $S = \langle U, AT, V, f\rangle$ be an IFIS, $B \subseteq AT, U = \{x_1, x_2, \ldots, x_n\}$. A dominance matrix of system S with respect to B is defined as

$$M_B = (m_{ij})_{n\times n} = \begin{cases} 1, x_j \in [x_j]_B^{\geqslant} \\ 0, \text{other} \end{cases} \; i, j = 1, 2, \ldots, n.$$

M_B is also called l level dominance matrix of S if $|B| = l$.

Definition 7. The intersection of the dominance matrices M_B and M_C is defined as follows: for any $B, C \subseteq AT$ on the $S = \langle U, AT, V, f\rangle$, $M_B \cap M_C = (m_{ij})_{n\times n} \cap (m'_{ij})_{n\times n} = (\min\{m_{ij}, m'_{ij}\})_{n\times n}$.

Property 3. Given $S = \langle U, AT, V, f\rangle$ and $B, C \subseteq AT$; if M_B, M_C are two dominance matrices, we have

1. $m_{ii} = 1, i = 1, 2, \ldots, n$;
2. if $B, C \subseteq AT$, then $M_{B\cup C} = M_B \cap M_C$.

Property 3 can be obtained directly from Definition 6 and 7.

Definition 8. Let $S = \langle U, AT, V, f\rangle$ be an IFIS, $B \subseteq AT$; a dominance matrix M_B with B. $|M_B|$ is the dominance cardinality of B if it indicated that the number of non-zero elements, that is, a total number of the value 1.

Theorem 6. Let $S = \langle U, AT, V, f\rangle$ be an IFIS; $B \subseteq AT$. Then $|M_B| = \sum_{i=1}^{n} |[x_j]_B^{\geqslant}|$, $GK(B) = \frac{1}{|U|^2}|M_B|$.

Proof. According to the definition of the dominance matrix, for $B \subseteq AT$, we have that $\{m_{i_1}, m_{i_2}, \ldots, m_{i_n}\}$ corresponds to $\{x'_{i_1}, x'_{i_2}, \ldots, x'_{i_n}\} = [x_j]_B^{\geqslant}$, where $x'_{i_k} = \begin{cases} x_j, \; x_j \in [x_j]_B^{\geqslant} \\ \emptyset, \text{ other} \end{cases}$. Therefore, $|M_B| = \sum_{i=1}^{n} |[x_j]_B^{\geqslant}|$. It holds that $GK(B) = \frac{1}{|U|^2}|M_B|$. This completes the proof. □

Definition 9. Given two n-dimensional $n \times 1$ vectors $\alpha = \{e_1, e_2, \ldots, e_n\}^T$ and $\beta = \{b_1, b_2, \ldots, b_n\}^T$ (T denotes transpose), α is smaller than β if $e_i \leq b_i$ $(i = 1, \ldots, n)$.

Definition 10. Let $M_A = \{\alpha_1, \alpha_2, \dots, \alpha_n\}^T$ and $M_B = \{\beta_1, \beta_2, \dots, \beta_n\}^T$ be a matrix, where α_i and β_i $(i = 1, \dots, n)$ are n-dimensional $n \times 1$ vectors. M_A is smaller than M_B if $\alpha_i \leq \beta_i$ $(i = 1, \dots, n)$, denoted by $M_A \leq M_B$.

4.2 Matrix Algorithm for Reduction of Knowledge

Definition 11. Let $S = \langle U, AT, V, f\rangle$ be an IFIS. $B \subseteq AT$ is a reduct of AT if $GK(B) = GK(AT)$. If there is not $b \in B$ such that $GK(B - \{b\}) = GK(AT)$, B is one of the maximum reducts about AT.

Let $S = \langle U, AT, V, f\rangle$ be an IFIS, $U = \{x_1, x_2, \dots, x_n\}$, $AT = \{a_1, a_2, \dots, a_m\}$, $B \subseteq AT$, $M_B = \{\beta_1, \beta_2, \dots, \beta_n\}^T$ and $M_{AT} = \{\gamma_1, \gamma_2, \dots, \gamma_n\}^T$. We design a greedy algorithm for reduction of knowledge based on the dominance matrix.

Matrix algorithm on knowledge reduction in IFIS

Input: IFIS $S = \langle U, AT, V, f\rangle$.

Output: One reduct B of AT.

Step 1. Compute the dominance matrix $M_{AT} = \{\gamma_1, \gamma_2, \dots, \gamma_n\}^T$ of AT.

Step 2. Compute the first level matrix for every $a_l \in AT (1 \leq l \leq m)$: $M_{\{a_l\}} = M^{(l)}_{\{a_l\}} = \{\tau^{(l)}_1, \tau^{(l)}_2, \dots, \tau^{(l)}_n\}^T$ of AT.

For $i = 1$ to n do

If $0 \neq \tau^{(l)}_i \leq \gamma_i$, then let $\tau^{(l)}_i = 0$, and the new matrix is denoted by $FM^{(l)}_{\{a_l\}}$, $FM^{(l)}_{\{a_i\}} = \{\tau^{(l)}_1, \tau^{(l)}_2, \dots, \tau^{(l)}_n\}^T$, $a_l \in AT$, $(1 \leq l \leq m)$ is called the first-level reduct matrix.

Come to the next step.

Step 3. If $FM^{(l)}_{\{a_i\}} = 0$ then output a first-level reduct $\{a_l\}$. Otherwise, go to the next step.

Step 4. All second-level dominance matrices are obtained by intersection of the non-0 first-level reduct matrices on step 2: $M^{(2)}_{\{a_l a_s\}}$, $M^{(2)}_{\{a_l a_s\}} \neq M^{(1)}_{\{a_l\}}$, $M^{(2)}_{\{a_l a_s\}} \neq M^{(1)}_{\{a_s\}}$, $l \neq s, l, s = 1, 2, \dots, n$.

Find all of the second-level reducts by the method used in Step 2.

Step 5. Repeat Step 4 to obtain to the third-level and more reducts, until $M^{(k)}_B = 0$ $(1 \leq k \leq m)$, $B \subseteq AT$.

The time complexity of this algorithm is $O(|U|^2 2^{|A|})$.

Example 7. Table 3 presents an IFIS $S = \langle U, AT, V, f\rangle$, where $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9\}$, $AT = \{a_1, a_2, a_3, a_4\}$. We demonstrate the effectiveness of our algorithms using it.

U	a_1	a_2	a_3	a_4
x_1	0.1	0.2	0.1	0.1
x_2	0.1	$*$	0.3	0.1
x_3	0.3	0.2	0.3	$*$
x_4	0.1	0.2	$*$	0.1
x_5	$*$	0.2	0.1	0.3
x_6	0.3	0.1	$*$	0.3
x_7	0.3	0.2	$*$	$*$
x_8	0.3	0.1	0.2	0.3
x_9	0.2	0.3	$*$	0.2

Table 3. An IFIS for Example 7

Step 1: Construct the dominance matrices.

$$
\begin{aligned}
M_{AT} &= \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix} \\
M_{\{a_1\}} &= \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1
\end{pmatrix} \\
M_{\{a_2\}} &= \begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{aligned}
$$

$$M_{\{a_3\}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$M_{\{a_4\}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Steps 2 and 3: Construct the first-level reduct matrices and output reducts. Compare their rows of $M_{\{a_1\}}$, $M_{\{a_2\}}$, $M_{\{a_3\}}$ and $M_{\{a_4\}}$ to M_{AT}. Find that there is no $0 \neq \tau_1^{(l)} \leq \gamma_i$. Values of row 1 and 4 are the same for $M_{\{a_2\}}$ and M_{AT}. Values of row 2 are the same for $M_{\{a_3\}}$ and M_{AT}. Values of row 6 and 8 are the same for $M_{\{a_4\}}$ and M_{AT}. Therefore, we can see that there is no first-level reduct.

Thus, the first-level reduction matrices are

$$FM_{\{a_1\}}^{\{1\}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$FM_{\{a_2\}}^{\{1\}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$FM^{\{1\}}_{\{a_3\}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$FM^{\{1\}}_{\{a_4\}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Step 4 and 5: Get the second-level and more dominance matrices, reducts.

$$M^{\{2\}}_{\{a_1,a_2\}} = FM^{\{1\}}_{\{a_1\}} \cap FM^{\{1\}}_{\{a_2\}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$M^{\{2\}}_{\{a_1,a_3\}} = FM^{\{1\}}_{\{a_1\}} \cap FM^{\{1\}}_{\{a_3\}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$M^{\{2\}}_{\{a_1,a_4\}} = FM^{\{1\}}_{\{a_1\}} \cap FM^{\{1\}}_{\{a_4\}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$M^{\{2\}}_{\{a_2,a_3\}} = FM^{\{1\}}_{\{a_2\}} \cap FM^{\{1\}}_{\{a_3\}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$M^{\{2\}}_{\{a_2,a_4\}} = FM^{\{1\}}_{\{a_2\}} \cap FM^{\{1\}}_{\{a_4\}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$M^{\{2\}}_{\{a_3,a_4\}} = FM^{\{1\}}_{\{a_3\}} \cap FM^{\{1\}}_{\{a_4\}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Compared with M_{AT}, there is no second-level reduct. Using the same method we can get $FM^{(2)}_{\{a_1,a_2,a_3\}}$. Values of each row of $FM^{(2)}_{\{a_1,a_3,a_4\}}$ are 0. The loop is stopped.

Therefore, we can see that a reduct of AT is $\{a_1, a_2, a_3\}$.

5 RESULTS OF EXPERIMENTS

In Table 4, results of experiments on six well-known data sets from the UCI Machine Learning Repository are cited [2]. The Matrix and IFSPA-IVPR [13] algorithm have been implemented using MATLAB for the databases. From the table, it is evident that Matrix algorithm produces reduct for large data sets with more attributes. The performance analysis of the Matrix and the IFSPA-IVPR is also depicted in Figure 1. (Note: at. = attributes, mv. = missing values, I-I. = IFSPA-IVPR, Mx. = Matrix)

Data set	Instances	No. of at.	No. of mv.	I-I.	Mx.
Car	8	4	5	2	2
Hepatitis	155	19	167	6	5
Heart (Switzerland)	123	13	273	5	6
Soybean (Large)	307	35	705	10	9
Water – treatment – data	527	38	592	7	7
Echocardiogram	74	13	132	2	2

Table 4. Comparative analysis of Matrix and IFSPA-IVPR algorithm

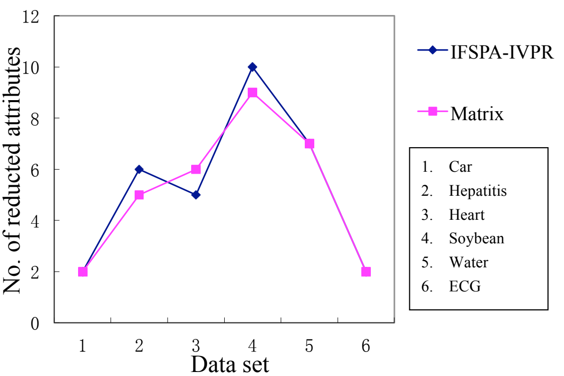

Figure 1. Performance analysis of the Matrix and the IFSPA-IVPR

6 CONCLUSIONS

Dynamic data, noise data and defect data, etc., make the analysis results instable and uncertain. It limits the application of rough set theory. In this paper, we define granulation of knowledge from the view of information, use rough entropy of attributes to define the significance of the attributes in IFIS, and discuss some important properties of them. From these properties, it can be shown that these measures which are proposed provide important approaches to measuring the discernibility ability of different knowledge in IFIS. As an application of the granulation and the rough entropy, we proposed definition of dominance matrix. A greedy algorithm based on the dominance matrix for knowledge reduction is proposed for finding the maximum reduct in IFIS. The time complexity of this algorithm is $O(|U|^2 2^{|A|})$. The importance of the maximum reduct is due to its potential for speeding up data process and improving the quality of classification. These new approaches may be helpful for rule evaluation and knowledge discovery in incomplete fuzzy information systems.

REFERENCES

[1] Beaubouef, T.—Petry, F. E.—Arora, G.: Information-Theoretic Measures of Uncertainty for Rough Sets and Rough Relational Databases. Information Sciences 1998, No. 109, pp. 185–195.

[2] Blake, C. L.—Merz, C. J.: UCI Repository of Machine Learning Databases. Irvine, University of California 1998. Availaible on: `http://www.ics.uci.edu/mlearn/`.

[3] Greco, S.—Matarazzo, B.—Slowinski, R.: Fuzzy Rough Sets and Multiple-Premise Gradual Decision Rules. International Journal of Approximate Reasoning, Vol. 41, 2006, No. 2, pp. 179–211.

[4] Greco, S.—Matarazzo, B.—Slowinski, R.: Rough Approximation by Dominance Relations. International Journal of Intelligent Systems, Vol. 17, 2002, No. 2, pp. 153–171.

[5] Greco, S.—Matarazzo, B.—Slowinski, R. S.: Rough Approximation of a Preference Relation by Dominance Relations. European Journal of Operational Research, Vol. 117, 1999, No. 1, pp. 63–83.

[6] Li, H. L.—Chen, M. H.: Induction of Multiple Criteria Optimal Classification Rules for Biological and Medical Data. Computers in Biology and Medicine, Vol. 38, 2008, No. 1, pp. 42–52.

[7] Liang, J. Y.—Xu, Z. B.: The Algorithm on Knowledge Reduction in Incomplete Information Systems. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 10, 2002, No. 2002, pp. 95–103.

[8] Liang, J. Y.—Shi, Z. Z.: The Information Entropy, Rough Entropy and Knowledge Granulation in Rough Set Theory. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 12, 2004, No. 1, pp. 37–46.

[9] Mieszkowiez, R. A.—Rolka, L.: Fuzziness in Information Systems. Electronic Notes in Theoretical Computer Science, Vol. 82, 2003, No. 4, pp. 164–173.

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

[11] Pawlak, Z.: Rough Sets and Intelligent Data Analysis. Information Sciences, Vol. 147, 2002, No. 1, pp. 1–12

[12] Pawlak, Z.—Skowron, A.: Rudiments of Rough Sets. Information Sciences, Vol. 177, 2007, No. 1, pp. 3–27.

[13] QIAN, Y. H.—LIANG, J. Y.—PEDRYCZ, W.—DANG, C. Y.: An Efficient Accelerator for Attribute Reduction from Incomplete Data in Rough Set Framework. Pattern Recognition, Vol. 44, 2011, pp. 1658–1670.

[14] SWINIARSKI, R. W.—SKOWRON, A.: Rough Set Method in Feature Selection and Recognition. Pattern Recognition Letters, Vol. 24, 2003, No. 6, pp. 833–849.

[15] SHEN, L. X.—LOH, H. T.: Applying Rough Sets to Market Timing Decisions. Decision Support Systems, Vol. 37, 2004, No. 4, pp. 583–597.

[16] SHAO, M. W.—ZHANG, W. X.: Dominance Relation and Rules in an Incomplete Ordered Information System. International Journal of Intelligent Systems, Vol. 20, 2005, pp. 13–27.

[17] WANG, X. X.—TSANG, E. C. C.—ZHAO, S. Y. et al.: Learning Fuzzy Rules from Fuzzy Samples Based on Rough Set Technique. Information Sciences, Vol. 177, 2007, No. 20, pp. 4493–4514.

[18] WEI, L. H.—TANG, Z. M.—YANG, X. B.—ZHANG, L. L.: Dominance-Based Rough Set Approach to Incomplete Fuzzy Information System. Proceedings of the 2008 IEEE International Conference on Granular Computing, GrC 2008, Hangzhou, China, August 2008, pp. 632–637.

[19] WU, W. Z.—ZHANG, W. X.—LI, H. Z.: Knowledge Acquisition in Incomplete Fuzzy Information Systems via the Rough Set Approach. Expert Systems, Vol. 20, 2003, No. 5, pp. 280–286.

[20] XU, W. H.—ZHANG, X. Y.—ZHANG, W. X.: Knowledge Granulation, Knowledge Entropy and Knowledge Uncertainty Measure in Ordered Information Systems. Applied Soft Computing, Vol. 9, 2009, pp. 1244–1251.

[21] YANG, X. B.—YANG, J. Y.—WU, C. et al.: Dominance-Based Rough Set Approach and Knowledge Reductions in Incomplete Ordered Information System. Information Sciences, Vol. 178, 2008, No. 4, pp. 1219–1234.

[22] YEE, L.—WU, W. Z.—ZHANG, W. X.: Knowledge Acquisition in Incomplete Information Systems: A Rough Set Approach. European Journal of Operational Research, Vol. 168, 2006, No. 1, pp. 164–180.

[23] YANG, X. B.—YU, D. J.—YANG, J. Y.—WEI, L. H.: Dominance-Based Rough Set Approach to Incomplete Interval-Valued Information System. Data and Knowledge Engineering, Vol. 68, 2009, No. 11, pp. 1331–1347.

[24] YANG, X. B.—WEI, L. H.—YU, D. J.—YANG, J. Y.: Rough Fuzzy Set in Incomplete Fuzzy Information System Based on Similarity Dominance Relation. Recent Patents on Computer Science, Vol. 2, 2009, pp. 68–74.

[25] ZHAI, L. Y.—KHOO, L. P.—ZHONG, Z. W.: A Dominance Based Rough set Approach to Kansei Engineering in Product Development. Expert Systems with Applications, Vol. 36, 2009, No. 1, pp. 393–402.

Qiang Wu is an Assistant Professor at the Department of Computer Science and Technology of Shaoxing University in Zhejing (China). He attained his Ph. D. in computer in 2008 at Shanghai University (China). His current research interests include formal concept analysis and rough set theory.