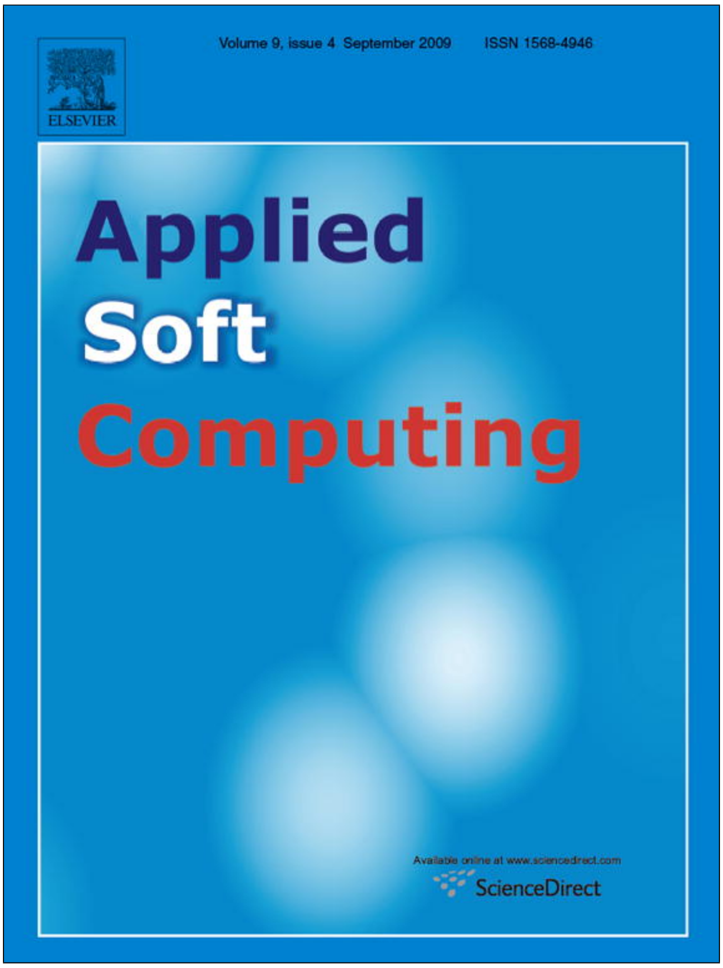
Volume 9, issue 4 September 2009 ISSN 1568-4946
ELSEVIER
Applied
Soft
Computing
Available online at www.sciencedirect.com
ScienceDirect

Applied Soft Computing 9 (2009) 1244–1251

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Knowledge granulation, knowledge entropy and knowledge uncertainty measure in ordered information systems

Xu Wei-hua [a,*], Zhang Xiao-yan [a], Zhang Wen-xiu [b]

[a] School of Mathematics and Physics, Chongqing Institute of Technology, Chongqing 400054, China
[b] School of Science, Xi'an Jiaotong University, Xi'an 710019, China

ARTICLE INFO

Article history:
Received 19 April 2007
Received in revised form 21 January 2009
Accepted 22 March 2009
Available online 1 April 2009

Keywords:
Rough set
Ordered information systems
Knowledge granulation
Knowledge entropy
Knowledge uncertainty measure

ABSTRACT

In this paper, concepts of knowledge granulation, knowledge entropy and knowledge uncertainty measure are given in ordered information systems, and some important properties of them are investigated. From these properties, it can be shown that these measures provides important approaches to measuring the discernibility ability of different knowledge in ordered information systems. And relationship between knowledge granulation, knowledge entropy and knowledge uncertainty measure are considered. As an application of knowledge granulation, we introduce definition of rough entropy of rough sets in ordered information systems. By an example, it is shown that the rough entropy of rough sets is more accurate than classical rough degree to measure the roughness of rough sets in ordered information systems.

1. Introduction

Rough set theory, proposed by Pawlak in the early 1980s [15], is an extension of the classical set theory for modeling uncertainty or imprecision information. The research has recently roused great interest in the theoretical and application fronts, such as machine learning, pattern recognition, data analysis, and so on.

In Pawlak's original rough set theory, partition or equivalence (indiscernibility relation) is an important and primitive concept. However, partition or equivalence relation, as the indiscernibility relation in Pawlak's original rough set theory, is still restrictive for many applications. To address this issue, several interesting and meaningful extensions to equivalence relation have been proposed in the past, such as tolerance relations [22], neighborhood operators [31], others [11,24,26–28,32]. Particularly, in many real situations, we are often face to the problems in which the ordering of properties of the considered attributes plays a crucial role. One such type of problem is the ordering of objects. For this reason, Greco, Matarazzo, and Slowinski proposed an extension rough set theory, called the dominance-based rough set approach (DRSA) to take into account the ordering properties of criteria [4–9]. This innovation is mainly based on substitution of the indiscernibility relation by a dominance relation. Moreover, Greco, Matarazzo, and Slowinski characterize the DRSA as well as decision rules induced from rough approximations, while the usefulness of the DRSA and its advantages over the CRSA (classical rough set approach) are presented [4–9]. In DRSA, condition attributes are criteria and classes are preference ordered. Several studies have been made about properties and algorithmic implementations of DRSA [2,3,23,29].

To evaluate uncertainty of a system, another important concept of entropy was introduced by Shannon in ref. [21]. It is a very useful mechanism for characterizing information contents in various modes and has been applied in diverse fields. The entropy and its variants were adapted for rough set theory in ref. [25] and information interpretation of rough set theory was given in refs. [16–18]. Beaubouef et al. [1] addressed information measures of uncertainty of rough sets and rough relation databases. In ref. [12], a new method for evaluating both uncertainty and fuzziness was proposed. Unlike most existing information entropies, Qian and Liang [19] proposed a so-called combination entropy for evaluating uncertainty of a knowledge from an information system. All these studies were dedicated to evaluating uncertainty of a set in terms of the partition ability of a knowledge. As a powerful mechanism, granulation was first introduced by Zadeh in ref. [33]. It presents a more visual and easily understandable description for a partition on the universe. To characterize the granulation, granular computing was introduced in ref. [34], which, as a term with many meanings, covers all the research related to granulations. With regard to

* Corresponding author.
E-mail addresses: datongxuweihua@126.com (X. Wei-hua), zxy19790915@163.com (Z. Xiao-yan), wxzhang@mail.xjtu.edu.cn (Z. Wen-xiu).

1568-4946/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.asoc.2009.03.007

granular computing, many pieces of nice work were accomplished in refs. [10,14,20,30]. Especially, closely associated with granular computing, several measures on knowledge in an information system were proposed and the relationships between these measures were discussed in ref. [13]. These measures include granulation measure, information entropy, rough entropy, and knowledge granulation, and have become effective mechanisms for evaluating uncertainty in rough set theory. In this paper, we introduce concepts of knowledge granulation, knowledge entropy and knowledge uncertainty measure in ordered information systems, and discuss some important properties of them. From these properties, it can be shown that these measures which are proposed provides important approaches to measuring the discernibility ability of different knowledge in ordered information systems.

The rest of this paper is organized as follows. Some preliminary concepts such as ordered information systems, indiscernibility relation, partition, lower and upper approximations, partial relation of knowledge and decision tables are briefly recalled in Section 2. In Sections 3–5, concepts of knowledge granulation, knowledge entropy and knowledge uncertainty measure in ordered information systems are introduced respectively, and some important properties of them are discussed. In Section 6, we investigate the relationship between knowledge granulation, knowledge entropy and knowledge uncertainty measure. Finally, as an application of knowledge granulation, we introduce definition of rough entropy of rough sets in ordered information systems in Section 7. By an example, it is shown that the rough entropy of rough sets is more accurate than classical rough degree to measure the roughness of rough sets in ordered information systems.

2. Rough sets and ordered information systems

The following recalls necessary concepts and preliminaries required in the sequel of our work. Detailed description of the theory can be found in the source papers [4–9]. A description has also been made in ref. [35].

The notion of information system (sometimes called data tables, attribute-value systems, knowledge representation systems, etc.) provides a convenient tool for the representation of objects in terms of their attribute values.

An information system is an ordered triple $\mathcal{I} = (U, A, F)$, where $U = \{x_1, x_2, \ldots, x_n\}$ is a non-empty finite set of objects called the universe, and $A = \{a_1, a_2, \ldots, a_p\}$ is a non-empty finite set of attributes, such that there exists a map $f_l : U \to V_{a_l}$ for any $a_l \in A$, where V_{a_l} is called the domain of the attribute a_l, and denoted $F = \{f_l | a_l \in A\}$.

In an information systems, if the domain of an attribute is ordered according to a decreasing or increasing preference, then the attribute is a criterion.

Definition 2.1. (See refs. [4–9]) An information system is called an ordered information system (OIS) if all condition attributes are criteria.

Assumed that the domain of a criterion $a \in A$ is complete pre-ordered by an outranking relation $\succcurlyeq_a$, then $x \succcurlyeq_a y$ means that x is at least as good as y with respect to criterion a. And we can say that x dominates y. In the following, without any loss of generality, we consider criterions having a numerical domain, that is, $V_a \subseteq \mathcal{R}$ ($\mathcal{R}$ denotes the set of real numbers).

We define $x \succcurlyeq y$ by $f(x, a) \geq f(y, a)$ according to increasing preference, where $a \in A$ and $x, y \in U$. For a subset of attributes $B \subseteq A$, $x \succcurlyeq_B y$ means that $x \succcurlyeq_a y$ for any $a \in B$, and that is to say x dominates y with respect to all attributes in B. Furthermore, we denote $x \succcurlyeq_B y$ by $x R_B^{\succcurlyeq} y$. In general, we denote a ordered information systems by $\mathcal{I}^{\succcurlyeq} = (U, A, F)$. Thus the following definition can be obtained.

Definition 2.2. (See refs. [4–9]) Let $\mathcal{I}^{\succcurlyeq} = (U, A, F)$ be an ordered information, for $B \subseteq A$, denote

$$R_B^{\succcurlyeq} = \{(x, y) \in U \times U | f_l(x) \geq f_l(y), \ \forall a_l \in B\};$$

$R_B^{\succcurlyeq}$ are called dominance relations of ordered information system $\mathcal{I}^{\succcurlyeq}$.

Let denote

$$[x_i]_B^{\succcurlyeq} = \{x_j \in U | (x_j, x_i) \in R_B^{\succcurlyeq}\} = \{x_j \in U | f_l(x_j) \geq f_l(x_i), \quad \forall a_l \in B\};$$
$$\frac{U}{R_B^{\succcurlyeq}} = \{[x_i]_B^{\succcurlyeq} | x_i \in U\},$$

where $i \in \{1, 2, \ldots, |U|\}$, then $[x_i]_B^{\succcurlyeq}$ will be called a dominance class or the granularity of information, and $U/R_B^{\succcurlyeq}$ be called a classification of U about attribute set B.

The following properties of a dominance relation are trivial by the above definition.

Proposition 2.1. (See refs. [4–9]) *Let $R_A^{\succcurlyeq}$ be a dominance relation.*

(1) $R_A^{\succcurlyeq}$ is reflexive, transitive, but not symmetric, so it is not an equivalence relation.
(2) If $B \subseteq A$, then $R_A^{\succcurlyeq} \subseteq R_B^{\succcurlyeq}$.
(3) If $B \subseteq A$, then $[x_i]_A^{\succcurlyeq} \subseteq [x_i]_B^{\succcurlyeq}$.
(4) If $x_j \in [x_i]_A^{\succcurlyeq}$, then $[x_j]_A^{\succcurlyeq} \subseteq [x_i]_A^{\succcurlyeq}$ and $[x_i]_A^{\succcurlyeq} = \cup\{[x_j]_A^{\succcurlyeq} | x_j \in [x_i]_A^{\succcurlyeq}\}$.
(5) $[x_j]_A^{\succcurlyeq} = [x_i]_A^{\succcurlyeq}$ iff $f(x_i, a) = f(x_j, a)$ for all $a \in A$.
(6) $|[x_i]_B^{\succcurlyeq}| \geq 1$ for any $x_i \in U$.
(7) $U/R_B^{\succcurlyeq}$ constitute a covering of U, i.e., for every $x \in U$ we have that $[x]_B^{\succcurlyeq} \neq \phi$ and $\bigcup_{x \in U} [x]_B^{\succcurlyeq} = U$.

where $|\cdot|$ denotes cardinality of the set.

For any subset X of U and A of $\mathcal{I}^{\succcurlyeq}$, the lower and upper approximation of X with respect to a dominance relation $R_A^{\succcurlyeq}$ could be defined as following (see refs. [4–9]):

$$\underline{R_A^{\succcurlyeq}}(X) = \{x \in U | [x]_A^{\succcurlyeq} \subseteq X\};$$
$$\overline{R_A^{\succcurlyeq}}(X) = \{x \in U | [x]_A^{\succcurlyeq} \cap X \neq \phi\}.$$

where $[x_i]_B^{\preccurlyeq} = \{x_j \in U | f_l(x_j) \leq f_l(x_i), \ \forall a_l \in B\}$.

From the above definition of rough approximation, the following important properties in ordered information systems have been proved, which are similar to those of Pawlak approximation spaces.

Proposition 2.2. (See refs. [4–9]) *Let $\mathcal{I}^{\succcurlyeq} = (U, A, F)$ be an ordered information system and $X \subseteq U$. The rough approximation can be expressed as union of elementary sets. That is to say the following holds.*

$$\underline{R_A^{\succcurlyeq}}(X) = \bigcup_{x \in U} \{[x]_A^{\succcurlyeq} | [x]_A^{\succcurlyeq} \subseteq X\};$$
$$\overline{R_A^{\succcurlyeq}}(X) = \bigcup_{x \in U} \{[x]_A^{\preccurlyeq} | [x]_A^{\succcurlyeq} \cap X \neq \phi\}.$$

Proposition 2.3. (See [4–9]) *Let $\mathcal{I}^{\succcurlyeq} = (U, A, F)$ be an ordered information system and $X, Y \subseteq U$, then its lower and upper approximations satisfy the following properties.*

$$\underline{R_A^{\succcurlyeq}}(X) \subseteq X \subseteq \overline{R_A^{\succcurlyeq}}(X). \tag{1}$$

$$\overline{R_A^{\succcurlyeq}}(X \cup Y) = \overline{R_A^{\succcurlyeq}}(X) \cup \overline{R_A^{\succcurlyeq}}(Y); \quad \underline{R_A^{\succcurlyeq}}(X \cap Y) = \underline{R_A^{\succcurlyeq}}(X) \cap \underline{R_A^{\succcurlyeq}}(Y). \tag{2}$$

$$\underline{R_A^{\succcurlyeq}}(X) \cup \underline{R_A^{\succcurlyeq}}(Y) \subseteq \underline{R_A^{\succcurlyeq}}(X \cup Y); \quad \overline{R_A^{\succcurlyeq}}(X \cap Y) \subseteq \overline{R_A^{\succcurlyeq}}(X) \cap \overline{R_A^{\succcurlyeq}}(Y). \tag{3}$$

$$\underline{R_X^{\succcurlyeq}}(\sim X) = \sim \overline{R_A^{\succcurlyeq}}(X); \quad \overline{R_A^{\succcurlyeq}}(\sim X) = \sim \underline{R_A^{\succcurlyeq}}(X). \tag{4}$$

$$\underline{R_A^{\succcurlyeq}}(U) = U; \quad \overline{R_A^{\succcurlyeq}}(\phi) = \phi. \tag{5}$$

$$\underline{R_A^{\succcurlyeq}}(X) = \underline{R_A^{\succcurlyeq}}(\underline{R_A^{\succcurlyeq}}(X)) = \overline{R_A^{\succcurlyeq}}(\underline{R_A^{\succcurlyeq}}(X)); \quad \overline{R_A^{\succcurlyeq}}(X) = \underline{R_A^{\succcurlyeq}}(\overline{R_A^{\succcurlyeq}}(X)) = \overline{R_A^{\succcurlyeq}}(\overline{R_A^{\succcurlyeq}}(X)). \tag{6}$$

$$\text{If } X \subseteq Y, \text{ then } \quad \underline{R_A^{\succcurlyeq}}(X) \subseteq \underline{R_A^{\succcurlyeq}}(Y) \text{ and } \overline{R_A^{\succcurlyeq}}(X) \subseteq \overline{R_A^{\succcurlyeq}}(Y). \tag{7}$$

where $\sim X$ is the complement of X.

Definition 2.3. For an ordered information system $\mathcal{I}^{\succcurlyeq} = (U, A, F)$ and $B, C \subseteq A$.

(1) If $[x]_B^{\succcurlyeq} = [x]_C^{\succcurlyeq}$ for any $x \in U$, then we call that classification $U/R_B^{\succcurlyeq}$ is equal to $R/R_C^{\succcurlyeq}$, denoted by $U/R_B^{\succcurlyeq} = U/R_C^{\succcurlyeq}$.
(2) If $[x]_B^{\succcurlyeq} \subseteq [x]_C^{\succcurlyeq}$ for any $x \in U$, then we call that classification $U/R_B^{\succcurlyeq}$ is finer than $R/R_C^{\succcurlyeq}$, denoted by $U/R_B^{\succcurlyeq} \subseteq U/R_C^{\succcurlyeq}$.
(3) If $[x]_B^{\succcurlyeq} \subseteq [x]_C^{\succcurlyeq}$ for any $x \in U$ and $[x]_B^{\succcurlyeq} \neq [x]_C^{\succcurlyeq}$ for some $x \in U$, then we call that classification $U/R_B^{\succcurlyeq}$ is properly finer then $R/R_C^{\succcurlyeq}$, denoted by $U/R_B^{\succcurlyeq} \subset U/R_C^{\succcurlyeq}$.

For an ordered information system $\mathcal{I}^{\succcurlyeq} = (U, A, F)$ and $B \subseteq A$, it is obtained that $U/R_A^{\succcurlyeq} \subseteq U/R_B^{\succcurlyeq}$ by Proposition 2.1(3) and above definition. So, an ordered information system $\mathcal{I}^{\succcurlyeq} = (U, A, F)$ be regarded as knowledge base $U/R_A^{\succcurlyeq}$, and $R_A^{\succcurlyeq}$ be regarded as knowledge.

Example 2.1. Given an ordered information system in Table 1.

From the table we can have

$[x_1]_A^{\succcurlyeq} = \{x_1, x_2, x_5, x_6\}$;

$[x_2]_A^{\succcurlyeq} = \{x_2, x_5, x_6\}$;

$[x_3]_A^{\succcurlyeq} = \{x_2, x_3, x_4, x_5, x_6\}$;

$[x_4]_A^{\succcurlyeq} = \{x_4, x_6\}$;

$[x_5]_A^{\succcurlyeq} = \{x_5\}$;

$[x_6]_A^{\succcurlyeq} = \{x_6\}$.

Table 1
An ordered information system.

U	a_1	a_2	a_3
x_1	1	2	1
x_2	3	2	2
x_3	1	1	2
x_4	2	1	3
x_5	3	3	2
x_6	3	2	3

If denote $B = \{a_1, a_2\}$, the following can be got

$[x_1]_B^{\succcurlyeq} = \{x_1, x_2, x_5, x_6\}$;

$[x_2]_B^{\succcurlyeq} = \{x_2, x_5, x_6\}$;

$[x_3]_B^{\succcurlyeq} = \{x_1, x_2, x_3, x_4, x_5, x_6\}$;

$[x_4]_B^{\succcurlyeq} = \{x_2, x_4, x_5, x_6\}$;

$[x_5]_B^{\succcurlyeq} = \{x_5\}$;

$[x_6]_B^{\succcurlyeq} = \{x_5, x_6\}$.

Thus, it is obviously that $U/R_A^{\succcurlyeq} \subseteq U/R_B^{\succcurlyeq}$. We can say that classification $U/R_A^{\succcurlyeq}$ is finer than classification $U/R_B^{\succcurlyeq}$, or knowledge $R_A^{\succcurlyeq}$ is finer than $R_B^{\succcurlyeq}$.

3. Knowledge granulation in ordered information systems

In this section, we will introduce a definition of granulation of knowledge in ordered information systems, and discuss some important properties.

Definition 3.1. Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, $R^{\geq}$ be a dominance relation, and $U/R^{\geq} = \{[u]_R^{\geq} \mid u \in U\}$ be the classification. Granulation of knowledge $R^{\geq}$, which is denoted by $GK(R^{\geq})$, is defined by

$$GK(R^{\geq}) = \frac{1}{|U|^2} \sum_{i=1}^{|U|} |[x_i]_R^{\geq}|.$$

Theorem 3.1. (Equivalence) *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $U/R^{\geq} = \{[u]_R^{\geq} \mid u \in U\}, U/S^{\geq} = \{[u]_S^{\geq} \mid u \in U\}$ be classifications of two dominance relations $R^{\geq}$ and $S^{\geq}$ respectively. If $|U/R^{\geq}| = |U/S^{\geq}|$, and it exists a bijective map $h : U/R^{\geq} \to U/S^{\geq}$ such that $|[u]_R^{\geq}| = |h([u]_R^{\geq})|$, then $GK(R^{\geq}) = GK(S^{\geq})$.*

Proof. It can be achieved by Definition 3.1. □

Corollary 3.1. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $R^{\geq}, S^{\geq}$ be two dominance relations. If $R^{\geq} = S^{\geq}$, then $GK(R^{\geq}) = GK(S^{\geq})$.*

Theorem 3.2. (Monotonicity) *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $U/R^{\geq} = \{[u]_R^{\geq} \mid u \in U\}, U/S^{\geq} = \{[u]_S^{\geq} \mid u \in U\}$ be classifications of two dominance relations $R^{\geq}$ and $S^{\geq}$ respectively. If $R^{\geq} \preccurlyeq S^{\geq}$, then $GK(R^{\geq}) \leq GK(S^{\geq})$.*

Proof. Because $R^{\geq} \preccurlyeq S^{\geq}$, it can be obtained $[u]_R^{\geq} \subseteq [u]_S^{\geq}$ for any $u \in U$. So, $|[u]_R^{\geq}| \leq |[u]_S^{\geq}|$. Thus, the following holds, i.e.,

$$GK(R^{\geq}) = \frac{1}{|U|^2} \sum_{i=1}^{|U|} |[x_i]_R^{\geq}| \leq \frac{1}{|U|^2} \sum_{i=1}^{|U|} |[x_i]_S^{\geq}| = GK(S^{\geq}).$$

Hence,

$GK(R^{\geq}) \leq GK(S^{\geq})$.

The theorem was proved. □

Example 3.1. (Continued from Example 2.1) By computing, we have that

$$GK(R_A^{\geq}) = \frac{1}{6^2}\cdot(4+3+5+2+1+1) = \frac{4}{9}$$
$$GK(R_B^{\geq}) = \frac{1}{6^2}\cdot(4+3+6+4+1+2) = \frac{5}{9}$$

Obviously, $GK(R_A^{\geq}) \leq GK(R_B^{\geq})$

By Theorem 3.2, we can acquire the following corollary.

Corollary 3.2. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $R^{\geq}, S^{\geq}$ be two dominance relations. If $R^{\geq} \prec S^{\geq}$, then $GK(R^{\geq}) < GK(S^{\geq})$.*

Corollary 3.3. Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $R^{\geq}, S^{\geq}$ be two dominance relations. If $R^{\geq} \preccurlyeq S^{\geq}$ and $GK(R^{\geq}) = GK(S^{\geq})$, then $R^{\geq} = S^{\geq}$.

Theorem 3.3. (Minimum) *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $R^{\geq}$ be a dominance relation. The minimum of knowledge granulation of this ordered information system is $1/|U|$. This value is achieved only if $R^{\geq} = I^{\geq}$, where $I^{\geq}$ is an unit dominance relation, i.e., $U/I^{\geq} = \{[u]_I^{\geq} = \{u\} | u \in U\}$.*

Proof. Since $U/I^{\geq} = \{[u]_I^{\geq} = \{u\} | u \in U\}$, so we have

$$GK(I^{\geq}) = \frac{1}{|U|^2}\sum_{i=1}^{|U|}|[x_i]_R^{\geq}| = \frac{1}{|U|^2}\sum_{i=1}^{|U|}1 = \frac{1}{|U|}.$$

Thus,

$$GK(I^{\geq}) = \frac{1}{|U|}.$$

The theorem was proved. □

Theorem 3.4. (Maximum) *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $R^{\geq}$ be a dominance relation. The maximum of knowledge granulation of this ordered information system is 1. This value is achieved only if $R^{\geq} = \delta^{\geq}$, where $\delta^{\geq}$ is an universe dominance relation, i.e., $U/\delta^{\geq} = \{[u]_R^{\geq} = U | u \in U\}$.*

Proof. Since $U/\delta^{\geq} = \{[u]_\delta^{\geq} = U | u \in U\}$, so we have

$$GK(\delta^{\geq}) = \frac{1}{|U|^2}\sum_{i=1}^{|U|}|[x_i]_R^{\geq}| = \frac{1}{|U|^2}\sum_{i=1}^{|U|}|U| = 1.$$

Thus,

$$GK(\delta^{\geq}) = 1.$$

The proof was completed. □

Theorem 3.5. (Boundedness) *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $R^{\geq}$ be a dominance relation, then knowledge granulation $GK(R^{\geq})$ exists the boundedness, i.e.,*

$$\frac{1}{|U|} \leq GK(R^{\geq}) \leq 1,$$

where $GK(R^{\geq}) = 1/|U|$ if and only if $R^{\geq} = I^{\geq}$, and $GK(R^{\geq}) = 1$ if and only if $R^{\geq} = \delta^{\geq}$.

Proof. It can be obtained by above theorems. □

Theorem 3.6. (Knowledge resolved) *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}$ be classification of dominance relation $R^{\geq}$. If some knowledge fragment $[u]_R^{\geq} (u \in U)$ can be resolved into two new knowledge fragments, and else fragments have no changes, where we denote new knowledge by $R'^{\geq}$, then $GK(R'^{\geq}) \leq GK(R^{\geq})$.*

Proof. Let Assume that $[x_i]_R^{\geq}$ of $U/R^{\geq}$ can be resolved into $[x_i]_{R'}^{\geq}$ and $[x_j]_{R'}^{\geq}$ $(i < j)$, where $[x_i]_R^{\geq} = [x_i]_{R'}^{\geq} \cup [x_j]_{R'}^{\geq}$, and $[x_i]_{R'}^{\geq} \subseteq [x_i]_R^{\geq}, [x_j]_{R'}^{\geq} \subseteq [x_j]_R^{\geq}$. So, we have

$$\frac{U}{R'^{\geq}} = \{[x_1]_R^{\geq}, [x_2]_R^{\geq}, \cdots, [x_i]_{R'}^{\geq}, \cdots, [x_j]_{R'}^{\geq}, \cdots, [x_{|U|}]_R^{\geq}\}.$$

That is to say

$$\begin{aligned} GK(R^{\geq}) &= \frac{1}{|U|^2}\sum_{t=1}^{|U|}|[x_t]_R^{\geq}| \\ &= \frac{1}{|U|^2}\sum_{t=1}^{i-1}|[x_t]_R^{\geq}| + \frac{1}{|U|^2}|[x_i]_R^{\geq}| + \frac{1}{|U|^2}\sum_{t=i+1}^{j-1}|[x_t]_R^{\geq}| \\ &\quad + \frac{1}{|U|^2}|[x_j]_R^{\geq}| + \frac{1}{|U|^2}\sum_{t=j+1}^{|U|}|[x_t]_R^{\geq}| \geq \frac{1}{|U|^2}\sum_{t=1}^{i-1}|[x_t]_R^{\geq}| \\ &\quad + \frac{1}{|U|^2}|[x_i]_{R'}^{\geq}| + \frac{1}{|U|^2}\sum_{t=i+1}^{j-1}|[x_t]_R^{\geq}| + \frac{1}{|U|^2}|[x_j]_{R'}^{\geq}| \\ &\quad + \frac{1}{|U|^2}\sum_{t=j+1}^{|U|}|[x_t]_R^{\geq}| \\ &= GK(R'^{\geq}) \end{aligned}$$

Thus,

$$GK(R'^{\geq}) \leq GK(R^{\geq}).$$

The theorem was proved. □

Corollary 3.4. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, $R^{\geq}$ be a dominance. If $R^{\geq}$ can be resolved into a new knowledge $R'^{\geq}$, then $GK(R'^{\geq}) \leq GK(R^{\geq})$.*

Theorem 3.7. (Knowledge composed) *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}$ be classification of dominance relation $R^{\geq}$. If a new knowledge fragment can be composed of two knowledge fragments of $R^{\geq}$, and else fragments have no changes, where we denote new knowledge by $R''^{\geq}$, then $GK(R^{\geq}) \leq GK(R''^{\geq})$.*

Proof. Let Assume that $[x_k]_{R''}^{\geq}$ can be composed of $[x_i]_R^{\geq}$ and $[x_j]_R^{\geq}$ of $U/R^{\geq}$ $(i, j < k)$, where $[x_k]_{R''}^{\geq} = [x_i]_R^{\geq} \cup [x_j]_R^{\geq}$, and $[x_k]_R^{\geq} \subseteq [x_k]_{R''}^{\geq}$. So we have

$$\frac{U}{R''^{\geq}} = \{[x_1]_R^{\geq}, [x_2]_R^{\geq}, \cdots, [x_i]_R^{\geq}, \cdots, [x_j]_R^{\geq}, \cdots, [x_k]_{R''}^{\geq}, \cdots, [x_{|U|}]_R^{\geq}\}.$$

Thus,

$$\begin{aligned} GK(R^{\geq}) &= \frac{1}{|U|^2}\sum_{t=1}^{|U|}|[x_t]_R^{\geq}| \\ &= \frac{1}{|U|^2}\sum_{t=1}^{k-1}|[x_t]_R^{\geq}| + \frac{1}{|U|^2}|[x_k]_R^{\geq}| + \frac{1}{|U|^2}\sum_{t=k+1}^{|U|}|[x_t]_R^{\geq}| \\ &\leq \frac{1}{|U|^2}\sum_{t=1}^{k-1}|[x_t]_R^{\geq}| + \frac{1}{|U|^2}|[x_k]_{R''}^{\geq}| + \frac{1}{|U|^2}\sum_{t=k+1}^{|U|}|[x_t]_R^{\geq}| = GK(R''^{\geq}) \end{aligned}$$

That is to say

$$GK(R^{\geq}) \leq GK(R''^{\geq}).$$

The theorem was proved. □

Corollary 3.5. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, $R^{\geq}$ be a dominance. If a new knowledge $R''^{\geq}$ can be composed of $R^{\geq}$, then $GK(R''^{\geq}) \leq GK(R^{\geq})$.*

From the above conclusions, it can be shown that a knowledge granulation provides an important approach to measuring the discernibility ability of a knowledge in ordered information systems. The smaller the knowledge granulation is, the stronger its discernibility ability is.

4. Knowledge entropy in ordered information systems

In this section, two definitions of knowledge rough entropy and knowledge information entropy will be proposed in ordered information systems, and some important properties were investigated.

4.1. Knowledge rough entropy in ordered information systems

Definition 4.1. Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, $R^{\geq}$ be a dominance relation, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}$ be the classification. Rough entropy of knowledge $R^{\geq}$, which is denoted by $E_r(R^{\geq})$, is defined by

$$E_r(R^{\geq}) = -\sum_{i=1}^{|U|} \frac{1}{|U|} \cdot \log_2 \frac{1}{|[x_i]_B^{\geq}|}.$$

Theorem 4.1. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}, U/S^{\geq} = \{[u]_S^{\geq} | u \in U\}$ be classifications of two dominance relations $R^{\geq}$ and $S^{\geq}$ respectively. We can have the following conclusions.*

(1) **(Equivalence)** If $|U/R^{\geq}| = |U/S^{\geq}|$, and it exists a bijective map $h : U/R^{\geq} \to U/S^{\geq}$ such that $|[u]_R^{\geq}| = |h([u]_R^{\geq})|$, then $E_r(R^{\geq}) = E_r(S^{\geq})$.
(2) **(Monotonicity)** If $R^{\geq} \preccurlyeq S^{\geq}$, then $E_r(R^{\geq}) \leq E_r(S^{\geq})$.
(3) **(Boundedness)** Information entropy of knowledge $R^{\geq}$ exists the boundedness, i.e.,

$$0 \leq E_r(R^{\geq}) \leq \log_2 |U|,$$

where $E_r(R^{\geq}) = 0$ if and only if $R^{\geq} = I^{\geq}$, and $E_r(R^{\geq}) = \log_2 |U|$ if and only if $R^{\geq} = \delta^{\geq}$.
(4) **(Knowledge resolved)** If $R^{\geq}$ can be resolved into a new knowledge $R'^{\geq}$, then $E_r(R'^{\geq}) \leq E_r(R^{\geq})$.
(5) **(Knowledge composed)** If a new knowledge $R''^{\geq}$ can be composed of $R^{\geq}$, then $E_r(R^{\geq}) \leq E_r(R''^{\geq})$.

Proof. The proof of them are similar to Theorems 3.1–3.7. □

Example 4.1. The following example shows that converse proposition of Theorem 4.1(2) does not hold.

Let denote $B' = \{a_1\}$ and $B'' = \{a_2\}$, so we have

$$[x_1]_{B'}^{\geq} = [x_3]_{B'}^{\geq} = \{x_1, x_2, x_3, x_4, x_5, x_6\};$$

$$[x_2]_{B'}^{\geq} = [x_5]_{B'}^{\geq} = [x_6]_{B'}^{\geq} = \{x_2, x_5, x_6\};$$

$$[x_4]_{B'}^{\geq} = \{x_2, x_4, x_5, x_6\},$$

and

$$[x_1]_{B''}^{\geq} = [x_2]_{B''}^{\geq} = [x_6]_{B''}^{\geq} = \{x_1, x_2, x_5, x_6\};$$

$$[x_3]_{B''}^{\geq} = [x_4]_{B''}^{\geq} = \{x_1, x_2, x_3, x_4, x_5, x_6\};$$

$$[x_5]_{B''}^{\geq} = \{x_5\}.$$

By computing, we can find $E_r(R_{B'}^{\geq}) = 1.98747$ and $E_r(R_{B''}^{\geq}) = 2.19499$. That is to say $E_r(R_{B'}^{\geq}) < E_r(R_{B''}^{\geq})$. But $R_{B'}^{\geq} \preccurlyeq R_{B''}^{\geq}$ does not hold.

Example 4.2. (Continued from Example 2.1) By computing, we have that

$$E_r(R_A^{\geq}) = \frac{1}{6}\cdot\log_2 4 + \frac{1}{6}\cdot\log_2 3 + \frac{1}{6}\cdot\log_2 5 + \frac{1}{6}\cdot\log_2 2 + \frac{1}{6}\cdot\log_2 1 + \frac{1}{6}\cdot\log_2 1 = 1.15115$$

$$E_r(R_B^{\geq}) = \frac{1}{6}\cdot\log_2 4 + \frac{1}{6}\cdot\log_2 3 + \frac{1}{6}\cdot\log_2 6 + \frac{1}{6}\cdot\log_2 4 + \frac{1}{6}\cdot\log_2 1 + \frac{1}{6}\cdot\log_2 2 = 1.52832$$

So, $E_r(R_A^{\geq}) \leq E_r(R_B^{\geq})$.

4.2. Knowledge information entropy in ordered information systems

Definition 4.2. Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, $R^{\geq}$ be a dominance relation, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}$ be the classification. Information entropy of knowledge $R^{\geq}$, which is denoted by $E(R^{\geq})$, is defined by

$$E(R^{\geq}) = \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|[x_i]_R^{\geq}|}{|U|}\right)$$

Theorem 4.8. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}, U/S^{\geq} = \{[u]_S^{\geq} | u \in U\}$ be classifications of two dominance relations $R^{\geq}$ and $S^{\geq}$ respectively. We can have the following conclusions.*

(1) **(Equivalence)** If $|U/R^{\geq}| = |U/S^{\geq}|$, and it exists a bijective map $h : U/R^{\geq} \to U/S^{\geq}$ such that $|[u]_R^{\geq}| = |h([u]_R^{\geq})|$, then $E(R^{\geq}) = E(S^{\geq})$.
(2) **(Monotonicity)** If $R^{\geq} \preccurlyeq S^{\geq}$, then $E(R^{\geq}) \geq E(S^{\geq})$.
(3) **(Boundedness)** Information entropy of knowledge $R^{\geq}$ exists the boundedness, i.e.,

$$0 \leq E(R^{\geq}) \leq 1 - \frac{1}{|U|},$$

where $E(R^{\geq}) = 1 - 1/|U|$ if and only if $R^{\geq} = I^{\geq}$, and $E(R^{\geq}) = 0$ if and only if $R^{\geq} = \delta^{\geq}$.
(4) **(Knowledge resolved)** If $R^{\geq}$ can be resolved into a new knowledge $R'^{\geq}$, then $E(R'^{\geq}) \geq E(R^{\geq})$.
(5) **(Knowledge composed)** If a new knowledge $R''^{\geq}$ can be composed of $R^{\geq}$, then $E(R^{\geq}) \geq E(R''^{\geq})$.

Proof. The proof of them are similar to Theorems 3.1–3.7. □

Example 4.3. (Continued from Example 2.1) By computing, we have that

$$E(R_A^{\geq}) = \frac{1}{6}\left(1 - \frac{4}{6}\right) + \frac{1}{6}\left(1 - \frac{3}{6}\right) + \frac{1}{6}\left(1 - \frac{5}{6}\right) + \frac{1}{6}\left(1 - \frac{2}{6}\right) + \frac{1}{6}\left(1 - \frac{1}{6}\right) + \frac{1}{6}\left(1 - \frac{1}{6}\right) = \frac{5}{9}$$

$$E(R_B^{\geq}) = \frac{1}{6}\left(1 - \frac{4}{6}\right) + \frac{1}{6}\left(1 - \frac{3}{6}\right) + \frac{1}{6}\left(1 - \frac{6}{6}\right) + \frac{1}{6}\left(1 - \frac{4}{6}\right) + \frac{1}{6}\left(1 - \frac{1}{6}\right) + \frac{1}{6}\left(1 - \frac{2}{6}\right) = \frac{4}{9}$$

Thus, we have $E(R_A^{\geq}) \geq E(R_B^{\geq})$.

5. Knowledge uncertainty measure in ordered information systems

In this section, another uncertainty measure will be introduced, which can provide another important approach to measuring the discernibility ability of a knowledge in ordered information systems.

Definition 5.1. Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, $R^{\geq}$ be a dominance relation, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}$ be the classification. Uncertainty measure of knowledge $R^{\geq}$, which is denoted by $G(R^{\geq})$, is defined by

$$G(R^{\geq}) = -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{|[x_i]_R^{\geq}|}{|U|}$$

Theorem 5.1. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}, U/S^{\geq} = \{[u]_S^{\geq} | u \in U\}$ be classifications of two dominance relations $R^{\geq}$ and $S^{\geq}$ respectively. We can have the following conclusions.*

(1) **(Equivalence)** If $|U/R^{\geq}| = |U/S^{\geq}|$, and it exists a bijective map $h : U/R^{\geq} \to U/S^{\geq}$ such that $|[u]_R^{\geq}| = |h([u]_R^{\geq})|$, then $G(R^{\geq}) = G(S^{\geq})$.
(2) **(Monotonicity)** If $R^{\geq} \preccurlyeq S^{\geq}$, then $G(R^{\geq}) \geq G(S^{\geq})$.
(3) **(Boundedness)** Uncertainty measure of knowledge $R^{\geq}$ exists the boundedness, i.e.,

$$0 \leq G(R^{\geq}) \leq \log_2 |U|,$$

where $G(R^{\geq}) = \log_2 |U|$ if and only if $R^{\geq} = I^{\geq}$, and $E(R^{\geq}) = 0$ if and only if $R^{\geq} = \delta^{\geq}$.
(4) **(Knowledge resolved)** If $R^{\geq}$ can be resolved into a new knowledge $R'^{\geq}$, then $G(R'^{\geq}) \geq G(R^{\geq})$.
(5) **(Knowledge composed)** If a new knowledge $R''^{\geq}$ can be composed of $R^{\geq}$, then $G(R^{\geq}) \geq G(R''^{\geq})$.

Proof. The proof of them are similar to Theorems 3.1–3.7. □

Example 5.1. (Continued from Example 2.1) By computing, we have that

$$G(R_A^{\geq}) = -\frac{1}{6}\cdot\log_2\frac{4}{6} - \frac{1}{6}\cdot\log_2\frac{3}{6} - \frac{1}{6}\cdot\log_2\frac{5}{6} - \frac{1}{6}\cdot\log_2\frac{2}{6} - \frac{1}{6}\cdot\log_2\frac{1}{6} - \frac{1}{6}\cdot\log_2\frac{1}{6} = 1.43381$$

$$G(R_B^{\geq}) = -\frac{1}{6}\cdot\log_2\frac{4}{6} - \frac{1}{6}\cdot\log_2\frac{3}{6} - \frac{1}{6}\cdot\log_2\frac{6}{6} - \frac{1}{6}\cdot\log_2\frac{4}{6} - \frac{1}{6}\cdot\log_2\frac{1}{6} - \frac{1}{6}\cdot\log_2\frac{2}{6} = 1.05664$$

Thus, we have $G(R_A^{\geq}) \geq G(R_B^{\geq})$.

6. Relationship between knowledge granulation, knowledge entropy and uncertainty measure

In this section, we will discuss relationship between knowledge granulation, knowledge entropy and uncertainty measure.

Theorem 6.1. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, $R^{\geq}$ be a dominance relation, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}$ be the classification. Relationship between knowledge granulation $GK(R^{\geq})$ and information entropy $E(R^{\geq})$ of knowledge $R^{\geq}$ is*

$$GK(R^{\geq}) + E(R^{\geq}) = 1.$$

Proof. Because of $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}$ is classification of dominance $R^{\geq}$, we can have

$$E(R^{\geq}) = \sum_{i=1}^{|U|} \frac{1}{|U|}\left(1 - \frac{|[x_i]_R^{\geq}|}{|U|}\right) = \sum_{i=1}^{|U|} \frac{1}{|U|} - \sum_{i=1}^{|U|^2} \frac{|[x_i]_R^{\geq}|}{|U|} = 1 - GK(R^{\geq})$$

i.e.,

$$GK(R^{\geq}) + E(R^{\geq}) = 1.$$

The proof was completed. □

Example 6.1. (Continued form Example 3.1 and 4.3) In Examples 3.1 and 4.3, we have acquired that

$$GK(R_A^{\geq}) = \frac{4}{9}; \quad GK(R_B^{\geq}) = \frac{5}{9};$$
$$E(R_A^{\geq}) = \frac{5}{9}; \quad E(R_B^{\geq}) = \frac{4}{9}.$$

So, the following is obvious.

$$GK(R_A^{\geq}) + E(R_A^{\geq}) = 1;$$
$$GK(R_B^{\geq}) + E(R_B^{\geq}) = 1.$$

Theorem 6.2. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, $R^{\geq}$ be a dominance relation, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}$ be the classification. Relationship between uncertainty measure $G(R^{\geq})$ and rough entropy $E_r(R^{\geq})$ of knowledge $R^{\geq}$ is*

$$G(R^{\geq}) + E_r(R^{\geq}) = \log_2 |U|.$$

Proof. Because of $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}$ is classification of dominance $R^{\geq}$, we can have

$$\begin{aligned} G(R^{\geq}) &= -\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{|[x_i]_R^{\geq}|}{|U|} = -\sum_{i=1}^{|U|} \frac{1}{|U|} (\log_2 |[x_i]_R^{\geq}| - \log_2 |U|) \\ &= -\left(-\sum_{i=1}^{|U|} \frac{1}{|U|} \log_2 \frac{1}{|[x_i]_R^{\geq}|}\right) + \log_2 |U| \sum_{i=1}^{|U|} \frac{1}{|U|} \\ &= -E_r(R^{\geq}) + \log_2 |U| \end{aligned}$$

i.e.,

$$G(R^{\geq}) + E_r(R^{\geq}) = \log_2 |U|$$

The proof was completed. □

Example 6.2. (Continued form Example 4.2 and 5.1) In Examples 4.2 and 5.1, we have acquired that

$$E_r(R_A^{\geq}) = 1.15115; \quad E_r(R_B^{\geq}) = 1.52832;$$
$$G(R_A^{\geq}) = 1.43381; \quad G(R_B^{\geq}) = 1.05664.$$

So, the following is obvious.

$$G(R_A^{\geq}) + E_r(R_A^{\geq}) = \log_2 |U|;$$
$$G(R_B^{\geq}) + E_r(R_B^{\geq}) = \log_2 |U|.$$

7. Application

7.1. Limitation of classical measures in ordered information systems

In this section, through two illustrative examples, we reveal the limitations of existing classical measures for evaluating uncertainty of a set and approximation accuracy of a rough classification in ordered information systems.

In refs. [4–9], authors proposed two numerical measures for evaluating uncertainty of a set: accuracy and roughness. The accuracy measure is equal to the degree of completeness of a knowledge about the given object set X, and is defined by the ratio of the cardinalities of the lower and upper approximation sets of X as follows:

$$\alpha_R(X) = \frac{|\underline{R^{\succcurlyeq}}(X)|}{|\overline{R^{\succcurlyeq}}(X)|}.$$

The roughness measure represents the degree of incompleteness of a knowledge about the set, and is calculated by subtracting the accuracy from one:

$$\rho_R(X) = 1 - \frac{|\underline{R^{\succcurlyeq}}(X)|}{|\overline{R^{\succcurlyeq}}(X)|}.$$

These measures take into account the number of elements in each of the approximation sets and are good metrics for evaluating uncertainty that arises from the boundary region. However, the accuracy and roughness do not provide the information that is caused by the uncertainty related to the granularity of the indiscernibility relation. Their limitations are revealed by the following example.

Example 7.1. (Continued form Example 2.1) In Example 2.1, we have known $U/R_A^{\succcurlyeq} \subseteq U/R_B^{\succcurlyeq}$, i.e., classification $U/R_A^{\succcurlyeq}$ is finer than classification $U/R_B^{\succcurlyeq}$ in the system.

For $X' = \{x_3, x_5, x_6\}$, we have

$$\underline{R_A^{\succcurlyeq}}(X') = \underline{R_B^{\succcurlyeq}}(X') = \{x_5, x_6\};$$
$$\overline{R_A^{\succcurlyeq}}(X') = \overline{R_B^{\succcurlyeq}}(X') = U.$$

Thus, by calculating, the rough degrees of X' about knowledge $R_B^{\succcurlyeq}$ and $R_A^{\succcurlyeq}$ can be obtained respectively, which are

$$\rho_A(X) = \rho_B(X) = \frac{2}{3}.$$

In other words, the uncertainty of knowledge $R_B^{\succcurlyeq}$ is larger than that of $R_A^{\succcurlyeq}$ in Example 2.1, but X' has the same rough degree. Therefore, it is necessary to find a new and more accurate uncertainty measure for rough sets in ordered information systems.

7.2. Rough entropy of rough sets in ordered information systems

In the next, concept of rough entropy will be proposed, and it will be shown that it is a new and more accurate uncertainty measure for rough sets in ordered information systems.

Definition 7.1. Let $\mathcal{I}^{\succcurlyeq} = (U, A, F)$ be an ordered information system and $B \subseteq A$. The rough entropy of a rough set $X \subseteq U$ about knowledge $R_B^{\succcurlyeq}$ is defined as follows:

$$E_{R_B^{\geq}}(X) = \rho_B(X) \cdot GK(R_B^{\succcurlyeq}).$$

Furthermore, the following property can be obtained about the entropy of rough sets.

Theorem 7.1. *Let $\mathcal{I}^{\geq} = (U, A, F)$ be an ordered information system, and $U/R^{\geq} = \{[u]_R^{\geq} | u \in U\}, U/S^{\geq} = \{[u]_S^{\geq} | u \in U\}$ be classifications of two dominance relations $R^{\geq}$ and $S^{\geq}$ respectively. For any $X \subseteq U$, we can have the following conclusions.*

(1) **(Equivalence)** If $|U/R^{\geq}| = |U/S^{\geq}|$, and it exists a bijective map $h : U/R^{\geq} \to U/S^{\geq}$ such that $|[u]_R^{\geq}| = |h([u]_R^{\geq})|$, then $E_{R^{\geq}}(X) = E_{R^{\geq}}(X)$.
(2) **(Monotonicity)** If $R^{\geq} \preccurlyeq S^{\geq}$, then $E_{R^{\geq}}(X) \leq E_{S^{\geq}}(X)$.
(3) **(Boundedness)** Rough entropy of rough set X about knowledge $R^{\geq}$ exists the boundedness, i.e.,

$$0 \leq E(R^{\geq}) \leq 1,$$

where $E(R^{\geq}) = 0$ if and only if $R^{\geq} = I^{\geq}$, and $E(R^{\geq}) = 1$ if and only if $R^{\geq} = \delta^{\geq}$.
(4) **(Knowledge resolved)** If $R^{\geq}$ can be resolved into a new knowledge $R'^{\geq}$, then $E_{R'^{\geq}}(X) \leq E_{R^{\geq}}(X)$.
(5) **(Knowledge composed)** If a new knowledge $R''^{\geq}$ can be composed of $R^{\geq}$, then $E_{R^{\geq}}(X) \leq E_{R''^{\geq}}(X)$.

Proof. The proof of them can be acquired directly by Theorems 3.1–3.7 and Definition 7.1 □

From the above, the rough entropy of rough sets is related not only to its own rough degree, but also to the uncertainty of knowledge in the ordered information systems.

Example 7.2. (Continued from Example 7.1) The rough entropy of X' in Example 7.1 is calculated about knowledge $R_B^{\succcurlyeq}$ and $R_A^{\succcurlyeq}$ respectively, which are

$$E_{R_B^{\geq}}(X') = \rho(X') \cdot GK(R_B^{\succcurlyeq}) = \frac{2}{3} \times \frac{5}{9} = \frac{10}{27};$$
$$E_{R_A^{\geq}}(X') = \rho(X') \cdot GK(R_A^{\succcurlyeq}) = \frac{2}{3} \times \frac{4}{9} = \frac{8}{27}$$

Thus, we have

$$E_{R_A^{\geq}}(X') < E_{R_B^{\geq}}(X').$$

By this example, it is obvious that the rough entropy of rough sets is more accurate than the rough degree to measure the roughness of rough sets in ordered information systems.

8. Conclusions

Rough set theory is a new mathematical tool to deal with vagueness and uncertainty. Development of a rough computational method is one of the most important research tasks. While, in practise, ordered information system confines the applications of classical rough set theory. In this article, we introduced concepts of knowledge granulation, knowledge entropy and knowledge uncertainty measure in ordered information systems, and discuss some important properties of them. From these properties, it can be shown that these measures which are proposed provides important approaches to measuring the discernibility ability of different knowledge in ordered information systems. As an application of knowledge granulation, we proposed definition of rough entropy of rough sets in ordered information systems. By an example, it is shown that the rough entropy of rough sets is more accurate than classical rough degree to measure the roughness of rough sets in ordered information systems. These new measures may be helpful for rule evaluation and knowledge discovery in ordered information systems.

References

[1] T. Beaubouef, F.E. Petry, G. Arora, Information-theoretic measures of uncertainty for rough sets and rough relational databases, Information Sciences 109 (1998) 185–195.
[2] K. Dembczynski, R. Pindur, R. Susmaga, Generation of exhaustive set of rules within Dominance-based rough set approach, Electronic Notes Theory Computer Science 82 (4) (2003).
[3] K. Dembczynski, R. Pindur, R. Susmaga, Dominance-based rough set classifier without induction of decision rules, Electronic Notes Theory Computer Science 82 (4) (2003).

[4] S. Greco, B. Matarazzo, R. Slowinski, Rough approximation of a preference relation by dominance relation, ICS Research Report 16/96, Warsaw University of Technology, 1996
S. Greco, B. Matarazzo, R. Slowinski, Rough approximation of a preference relation by dominance relation, Europe Journal of Operation Research 117 (1999) 63–83.
[5] S. Greco, B. Matarazzo, R. Slowinski, A new rough set approach to multicriteria and multiattribute classification, in: L. Polkowsik, A. Skowron (Eds.), Rough Sets and Current Trends in Computing (RSCTC'98), Lecture Notes in Artificial Intelligence, vol 1424., Springer-Verlag, Berlin, 1998, pp. 60–67.
[6] S. Greco, B. Matarazzo, R. Slowinski, A new rough sets approach to evaluation of bankruptcy risk, in: X. Zopounidis (Ed.), Operational tools in the management of financial risks, Kluwer, Dordrecht, 1999, pp. 121–136.
[7] S. Greco, B. Matarazzo, R. Slowinski, Rough set theory for multicriteria decision analysis, Europe Journal of Operation Research 129 (2001) 11–47.
[8] S. Greco, B. Matarazzo, R. Slowinski, Rough sets methodology for sorting problems in presence of multiple attributes and criteria, European Journal of Operation Research 138 (2002) 247–259.
[9] S. Greco, B. Matarazzo, R. Slowinski, Dominance-based Rough Set approach as a proper way of handling graduality in rough set theory, Transaction on Rough Sets VII, Lecture Notes in Computer Science 4400 (2007) 36–52.
[10] G.J. Klir, Basic issues of computing with granular computing, in: Proceedings of IEEE International Conference on Fuzzy Systems, 1998, pp. 101–105.
[11] Y. Leuang, W.Z. Wu, W.X. Zhang, Knowledge acquisition in incomplete information systems: a rough set approach, European Journal of Operational Research 168 (1) (2006) 164–180.
[12] J.Y. Liang, K.S. Chin, C.Y. Dang, C.M. YAM, A new method for measuring uncertainty and fuzziness in rough set theory, International Journal of General Systems 31 (4) (2002) 331–342.
[13] J.Y. Liang, Z.Z. Shi, D.Y. Li, Information entropy, rough entropy and knowledge granulation in incomplete information systems, International Journal of General Systems 35 (6) (2006) 641–654.
[14] T.Y. Lin, Granular computing on binary relations I: Data mining and neighborhood systems, II: Rough sets representations and belief functions, in: L. Polkowski, A. Skowron (Eds.), Rough Sets in Knowledge Discovery, vol. 1, Physica-Verlag, Heidelberg, 1998, pp. 107–140.
[15] Z. Pawlak, Rough sets, International Journal of Computer and Information Science 11 (1982) 341–356.
[16] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991.
[17] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (2007) 3–27.
[18] Z. Pawlak, A. Skowron, Rough sets: some extensions, Information Sciences 177 (2007) 28–40.
[19] Y.H. Qian, J.Y. Liang, Combination entropy and combination granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 16 (2) (2008) 179–193.
[20] Y.H. Qian, J.Y. Liang, D.Y. Li, H.Y. Zhang, C.Y. Dang, Measures for evaluating the decision performance of a decision table in rough set theory, Information Sciences 178 (1) (2008) 181–202.
[21] C.E. Shannon, The mathematical theory of communication, The Bell System Technical Journal 27 (3,4) (1948) 373–423.
[22] A. Skowron, J. Stepaniuk, Tolerance approxximation space, Fundamental Information 27 (1996) 245–253.
[23] R. Susmaga, R. Slowinski, S. Greco, B. Matarazzo, Generation of reducts and rules in multi-attribute and multi-criteria classification, Control and Cybernetics 4 (2000) 969–988.
[24] K. Thangavel, A. Pethalakshmi, Dimensionality reduction based on rough set theory: A review, Applied Soft Computing 9 (2009) 1–12.
[25] M.J. Wierman, Measuring uncertainty in rough set theory, International Journal of General Systems 28 (1999) 283–297.
[26] W.Z. Wu, Y. Leung, W.X. Zhang, Connections between rough set theory and Dempster-Shafer theory of evidence, International Journal of General Systems 31 (2002) 405–430.
[27] W.Z. Wu, M. Zhang, H.Z. Li, J.S. Mi, Knowledge reduction in random information systems via Dempster-Shafer theory of evidence, Information Sciences 174 (2005) 143–164.
[28] W.H. Xu, W.X. Zhang, Measuring roughness of generalized rough sets induced by a covering, Fuzzy Sets and Systems 158 (2007) 2443–2455.
[29] W.H. Xu, M.W. Shao, Zhang Knowledge reduction based on evidence reasoning theory in ordered information systems, Lecture Notes in Artificial Intelligence (2006) 535–547.
[30] Y.Y. Yao, Granular computing on basic issues and possible solutions, in: Proceedings of the Fifth International Conference on Computing and Information, vol. I, 2000, pp. 186–189.
[31] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Information, Sciences 101 (1998) 239–259.
[32] D.R. Yu, Q.H. Hu, C.X. Wu, Uncertainty measures for fuzzy relations and their applications, Applied Soft Computing 7 (2007) 1135–1143.
[33] L.A. Zadeh, Fuzzy sets and information granularity, in: M. Gupta, R. Ragade, R. Yager (Eds.), Advances in Fuzzy Set Theory and Application, North-Holland, Amsterdam, 1979, pp. 3–18.
[34] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90 (1997) 111–127.
[35] W.X. Zhang, W.Z. Wu, J.Y. Liang, D.Y. Li, Theory and Method of Rough Sets, Science Press, Beijing, 2001.