Information Sciences 232 (2013) 143–166

Contents lists available at SciVerse ScienceDirect

Information Sciences

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

Can fuzzy entropies be effective measures for evaluating the roughness of a rough set?

Wei Wei [a], Jiye Liang [a,*], Yuhua Qian [a], Chuangyin Dang [b]

[a] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China

[b] Department of System Engineering and Engineering Management, City University of Hong Kong, Hong Kong

ARTICLE INFO

Article history:
Received 5 May 2011
Received in revised form 22 December 2012
Accepted 27 December 2012
Available online 8 January 2013

Keywords:
Rough set
Fuzzy set
Roughness
Fuzzy entropy
Granular computing

ABSTRACT

The roughness of a rough set arises from the existence of its boundary region. In such a boundary region, each object has a non-zero rough membership degree. When an object's rough membership degree is regarded as its fuzzy membership degree, a rough set can induce a fuzzy set. This relationship motivates us to assert that there may exist some inherent relations between the roughness of a rough set and the fuzziness of the fuzzy set induced from the rough set. This assertion leads us to the question: Can the existing fuzzy entropies be used to evaluate the roughness of a rough set? To answer this question, we first analyze how the boundary region varies when the partition of the universe becomes coarser, and then exploit this analysis in the introduction of a more appropriate definition on the roughness of a rough set. To determine whether a fuzzy entropy can be used to evaluate the roughness of a rough set or not, we develop three methods for estimating the ability of a fuzzy entropy to measure the roughness. The experiments show that these methods are very effective and can be applied to select a fuzzy entropy as a measure of the roughness of a rough set.

1. Introduction

Proposed by Pawlak in [36], rough set theory is based on the assumption that every object of the universe is associated with certain information (data, knowledge). The main goal of rough set theory is to synthesize the approximation of concepts from the acquired data [19,39]. Fuzzy set theory was introduced by Zadeh in [58], which provides an effective tool for representing vague concepts by allowing partial memberships [53]. It addresses the ill-defined boundary of a class through a continuous generalization of set characteristic functions. As a generalization of classical set theory, both rough set theory and fuzzy set theory have been used to model uncertainty [16]. As pointed out in [59], a fuzzy set characterizes the uncertainty that results from a class with unsharp boundaries, whereas a rough set describes the uncertainty generated from coarsely describing a crisp set.

On the connections and differences between rough set theory and fuzzy set theory is a fundamental question [37]. There have been many studies on this topic. Most researchers generally accept that the two theories are related but distinct [1,6,16,30,37]. Therefore, it is very significant to integrate the two theories in terms of the construction of models and measures of uncertainty. To date, many relevant papers have been published in the literature. First, we review several represen-

* Corresponding author. Tel./fax: +86 0351 7018176.
E-mail addresses: weiwei@sxu.edu.cn (W. Wei), ljy@sxu.edu.cn (J. Liang), jinchengqyh@126.com (Y. Qian), mecdang@cityu.edu.hk (C. Dang).

0020-0255/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2012.12.036

tative studies on the construction of models. Dubois and Prade [7] combined rough set theory with fuzzy set theory to define the fuzzy rough sets. This model employed the min and max fuzzy operators to describe the fuzzy lower and upper approximations. Radzikowska and Kerre [43] defined fuzzy rough sets in a more general manner based on the *T*-equivalence relation. The fuzzy lower and upper approximations were constructed by an implicator and a triangular norm. Mi and Zhang [29] presented a new fuzzy rough set definition based on a residual implication θ and its dual σ. Yeung et al. [57] defined lower and upper approximations based on arbitrary fuzzy relations using a constructive approach. In the following, several representative studies on uncertainty measures of a rough set will be reviewed. Chakrabarty et al. [5] proposed a fuzziness measure of a rough set, which is essentially a fuzzy entropy of the fuzzy set induced by the rough set's boundary region. Xu and Zhang [47] presented new lower and upper approximations in a generalized rough set induced by a covering and provided a measure of roughness by a rough membership function in covering approximation spaces. Yang and John [51] investigated the propagation of the roughness of rough sets under union, intersection, difference and complement operations, and determined their bounds under these situations. Qian et al. [40,41] defined fuzziness measures to compute the fuzziness of classification. Hu et al. [11–13] introduced probability into fuzzy approximation space and presented a theory of fuzzy probabilistic approximation spaces that combined probability, fuzziness and roughness into a rough set model. These studies are helpful for establishing the relationship between rough and fuzzy sets and understanding the essence of the uncertainty in decision tables.

Based on some of the studies mentioned above, one can find that by means of a rough membership function [19,39], a rough set determines one corresponding fuzzy set, and the fuzziness of the rough set can be evaluated by fuzzy entropies. To date, many fuzzy entropies have been proposed for evaluating fuzziness. Pal and Bezdek [32,33] surveyed several fuzzy entropies for a finite universal set. Liu [24] suggested an axiomatic definition of entropy, distance measure and similarity measure as well as the relationship between these three concepts. Kosko [18] observed the relationship between distance measure and fuzzy entropy from another viewpoint. Using Kosko's subsethood measure, Bhandari and Pal [3] provided a fuzzy information measure to estimate the discrimination between two fuzzy sets, defined an information distance known as divergence between fuzzy sets, and deduced a new fuzzy entropy. Luca and Termini [25] defined a fuzzy logarithm entropy formula on a limit universe. Pal and Pal [31] analyzed the classical Shannon's information entropy and proposed a new fuzzy entropy known as exponential entropy which can be applied to image-processing. Fan and Xie [8,9] proposed a method to generate a fuzzy entropy by distance measure and described a fuzzy entropy generated by different σ-similarity measures. All of the above fuzzy entropies of fuzzy sets provide effective methods for measuring the fuzziness of a fuzzy set.

For convenience of our further discussion, we also review several representative measures for evaluating the roughness of a rough set. Pawlak [35,36,38] defined a roughness measure that reflects the ratio of the number of objects in a rough set's lower approximation to that in its upper approximation. Bianucci and Cattaneo [4] noted that the roughness measure is not strictly monotonic with partition ordering. To solve this problem, Beaubouef et al. [2] investigated the measure of uncertainty based on information theory in rough sets and rough relational databases by proposing a rough entropy, which is significantly better than the Pawlak's roughness measure. Liang et al. [20–23] introduced the concepts of information entropy, rough entropy and knowledge granulation in rough set theory, in which the measures of accuracy, roughness and approximation accuracy are enhanced. Studies have been performed on specific applications, such as image segmentation and clustering. Pal et al. [34] used the idea of image granules to define the rough entropy of an image and employed the entropy in the processing of image segmentation. Małyszko and Stepaniuk [26,27] proposed new algorithmic schemes for standard and fuzzy rough entropy clustering algorithms (RECA) in rough entropy-based partitioning routines. Małyszko and Stepaniuk [28] designed a new algorithm based on granular multilevel rough entropy evolutionary thresholding (MRET) that operates on a multilevel domain.

Pawlak's roughness measure depends on the size of the boundary region and upper approximation, which results in the roughness measure not possessing strict monotonicity with partition ordering. Rough entropy depends on the size and structure of the boundary region, the positive region and the negative region. In fact, the roughness of a rough set is caused by those objects lying in the boundary region of a rough set. Therefore, there are certain shortcomings in the existing roughness measures. To overcome these shortcomings, we attempt to introduce a more appropriate measure for evaluating the roughness of a rough set. In the analysis, we observe that the roughness of a rough set arises from its boundary region and that the rough membership degree of each object in the rough set's boundary region is not zero. When an object's rough membership degree is regarded as its fuzzy membership degree, a rough set can induce a fuzzy set. This leads to the question: can the fuzzy entropies of a fuzzy set be used to evaluate the roughness of a rough set? To answer this question in this paper, we first review several common fuzzy entropies. We then develop three methods (JM1–JM3) for determining whether those entropies can be used to evaluate the roughness of a rough set. The experiments show that these methods provide an effective way to select an entropy as an appropriate roughness measure.

The remainder of the paper is organized as follows. Section 2 reviews some preliminary concepts and illustrates our motivation. Section 3 presents three methods of judging whether fuzzy entropies can evaluate the roughness of a rough set. Section 4 analyzes several common fuzzy entropies using the proposed methods. Section 5 demonstrates effectiveness of the proposed methods by experiments. Section 6 concludes the paper with some remarks.

2. Preliminaries

2.1. Rough set

An information system (also known as a data table, an attribute-value system, a knowledge representation system) is a basic concept in rough set theory and provides a convenient framework for the representation of objects in terms of their attribute values. Let $S=(U,A)$ be an information system, where U is a non-empty and finite set of objects, known as a universe, and A is a non-empty and finite set of attributes. For each $a\in A$, a mapping $a: U\rightarrow V_a$ is determined by an information system, where V_a is the domain of a.

Each non-empty subset $B\subseteq A$ determines an indiscernibility relation in the following way: $R_B=\{(x,y)\in U\times U|a(x)=a(y),\forall a\in B\}$, where $a(x)$ and $a(y)$ denote the values of objects x and y under a condition attribute a, respectively.

The relation R_B partitions U into equivalence classes given by $U/R_B=\{[x]_B|x\in U\}$, just U/B, where $[x]_B$ denotes the equivalence class determined by x with respect to B, i.e., $[x]_B=\{y\in U|(x,y)\in R_B\}$. Furthermore, for any $Y\subseteq U$, one defines that $(\overline{B}(Y),\underline{B}(Y))$ is the rough set of Y with respect to B, where the lower approximation $\underline{B}(Y)$ and the upper approximation $\overline{B}(Y)$ of Y [36,38,54] are described by

$$\underline{B}(Y)=\{x|[x]_B\subseteq Y\}\quad \text{and } \overline{B}(Y)=\{x|[x]_B\cap Y\neq\emptyset\}.$$

The objects in $\underline{B}(Y)$ are with certainty classified as members of Y on the basis of knowledge in B, while the objects in $\overline{B}(Y)$ are only classified as possible members of Y on the basis of knowledge in B. The set $BN_B(Y)=\overline{B}(Y)-\underline{B}(Y)$ is called the B-boundary region of Y, and consists of those objects that cannot be decisively classified based on the knowledge in B. A set is said to be rough (or crisp) if its boundary region is non-empty (or empty) [19].

Furthermore, we define a partial relation $\preceq$ on the family $\{U/B|B\subseteq A\}$ as follows: $U/A\preceq U/B$ if and only if for every $[x_i]_A$ there exists $[x_j]_B$ such that $[x_i]_A\subseteq[x_j]_B,\ x_i,x_j\in U$. In this case, we say that B is coarser than A (or A is finer than B). If $U/A\preceq U/B$ and $U/A\neq U/B$, we say that B is strictly coarser than A (or A is strictly finer than B), denoted by $U/A\prec U/B$ (or $U/B\succ U/A$).

2.2. Fuzzy entropy

In this paper, $\mathbb{R}^+=[0,+\infty)$, $\mathbf{F}(U)$ is the class of all fuzzy sets of U, and $\mathbf{P}(U)$ is the set of all of the crisp sets on the universe U. $\widetilde{X_1}=\frac{\mu_{\widetilde{X_1}(x_1)}}{x_1}+\frac{\mu_{\widetilde{X_1}(x_2)}}{x_2}+\cdots+\frac{\mu_{\widetilde{X_1}(x_n)}}{x_n}$ is a fuzzy set, where $\mu_{\widetilde{X_1}}(x)$ is the membership function of $\widetilde{X_1}\in\mathbf{F}(U)$. $\widetilde{X_1}^c$ indicates the complement of $\widetilde{X_1}$, i.e. $\mu_{\widetilde{X_1}^c}(x)=1-\mu_{\widetilde{X_1}}(x)$. For two fuzzy sets $\widetilde{X_1},\widetilde{X_2}\in\mathbf{F}(U)$, $\widetilde{X_1}\cup\widetilde{X_2}$ (the union of $\widetilde{X_1}$ and $\widetilde{X_2}$) is defined as $\mu_{\widetilde{X_1}\cup\widetilde{X_2}}(x)=\max\{\mu_{\widetilde{X_1}(x)},\mu_{\widetilde{X_2}(x)}\}$, and $\widetilde{X_1}\cap\widetilde{X_2}$ (the intersection of $\widetilde{X_1}$ and $\widetilde{X_2}$) is defined as $\mu_{\widetilde{X_1}\cap\widetilde{X_2}}(x)=\min\{\mu_{\widetilde{X_1}(x)},\mu_{\widetilde{X_2}}(x)\}$. The fuzzy set $\widetilde{X_1}^*$ is known as the sharpness of $\widetilde{X_1}$ if $\mu_{\widetilde{X_1}^*}(x)\geqslant\mu_{\widetilde{X_1}}(x)$ when $\mu_{\widetilde{X_1}}(x)>\frac{1}{2}$ and $\mu_{\widetilde{X_1}^*}(u)\leqslant\mu_{\widetilde{X_1}}(x)$ when $\mu_{\widetilde{X_1}}(x)\leqslant\frac{1}{2}$. For fuzzy set $\widetilde{X_1}$, the crisp sets $\widetilde{X}_{1\,near},\widetilde{X}_{1\,far}\in\mathbf{P}(U)$ are defined as

$$\mu_{\widetilde{X}_{1\,near}}(x)=\begin{cases}0, & \mu_{X_1}(x)\leqslant\frac{1}{2}\\ 1, & \mu_{X_1}(x)>\frac{1}{2}\end{cases},\quad \mu_{\widetilde{X}_{1\,far}}(x)=\begin{cases}0, & \mu_{X_1}(x)\geqslant\frac{1}{2}\\ 1, & \mu_{X_1}(x)<\frac{1}{2}\end{cases}.$$

Definition 2.1 [8,10]. A real function $d:\mathbf{F}^2(U)\rightarrow\mathbb{R}^+$ is a distance measure if d satisfies the following properties:

(DP1) $d(\widetilde{X_1},\widetilde{X_2})=d(\widetilde{X_2},\widetilde{X_1})$, $\forall\widetilde{X_1},\widetilde{X_2}\in\mathbf{F}(U)$;

(DP2) $d(\widetilde{X_1},\widetilde{X_1})=0,\ \forall\widetilde{X_1}\in\mathbf{F}(U)$;

(DP3) $d(D,D^c)=\max_{\widetilde{X_1},\widetilde{X_2}\in\mathbf{F}(U)}d(\widetilde{X_1},\widetilde{X_2}),\ \forall D\in\mathbf{P}(U)$;

(DP4) $\forall\widetilde{X_1},\widetilde{X_2},\widetilde{X_3}\in\mathbf{F}(U)$, if $\widetilde{X_1}\subset\widetilde{X_2}\subset\widetilde{X_3}$, then $d(\widetilde{X_1},\widetilde{X_2})\leqslant d(\widetilde{X_1},\widetilde{X_3})$ and $d(\widetilde{X_2},\widetilde{X_3})\leqslant d(\widetilde{X_1},\widetilde{X_3})$.

Let $U=\{x_1,x_2,\ldots,x_n\}$ be a finite set, $\widetilde{X_1},\widetilde{X_2}\in\mathbf{F}(U)$ be two fuzzy sets. The fuzzy Minkowski distance D_M^p between $\widetilde{X_1}$ and $\widetilde{X_2}$ is defined as [17]

$$D_M^p(\widetilde{X_1},\widetilde{X_2})=\left(\sum_{i=1}^{n}|\mu_{\widetilde{X_1}}(x_i)-\mu_{\widetilde{X_2}}(x_i)|^p\right)^{1/p},\quad p\geqslant 1,$$

where $\mu_{\widetilde{X_1}}(x)$ and $\mu_{\widetilde{X_2}}(x)$ are the membership functions of x in $\widetilde{X_1},\widetilde{X_2}\in\mathbf{F}(U)$, respectively.

Furthermore, Liu [24] noted that a distance is a σ-distance measure on $\mathbf{F}(U)$ if for any $\widetilde{X_1},\widetilde{X_2}\in\mathbf{F}(U)$ and $D\in\mathbf{P}(U)$, $d(\widetilde{X_1},\widetilde{X_2})=d(\widetilde{X_1}\cap D,\widetilde{X_2}\cap D)+d(\widetilde{X_1}\cap D^c,\widetilde{X_2}\cap D^c)$ holds.

Klir and Fogler [17] indicated that the fuzzy Hamming distance $D_H(\widetilde{X_1},\widetilde{X_2})=D_M^1(\widetilde{X_1},\widetilde{X_2})=\sum_{i=1}^{n}|\mu_{\widetilde{X_1}}(x_i)-\mu_{\widetilde{X_2}}(x_i)|$ is a σ-distance.

Bhandari and Pal [3] introduced a σ-distance that satisfies the conditions in [24] as follows:

$$D_L(\widetilde{X_1},\widetilde{X_2})=\sum_{i=1}^{n}\left((\mu_{\widetilde{X_1}}(x)-\mu_{\widetilde{X_2}}(x))\ln\frac{1+\mu_{\widetilde{X_1}}(x)}{1+\mu_{\widetilde{X_2}}(x)}+(\mu_{\widetilde{X_2}}(x)-\mu_{\widetilde{X_1}}(x))\ln\frac{2-\mu_{\widetilde{X_1}}(x)}{2-\mu_{\widetilde{X_2}}(x)}\right),$$

Fan et al. [8] defined another σ-distance as

$$D_E(\widetilde{X_1},\widetilde{X_2})=\sum_{i=1}^{n}\left(2-(1-\mu_{\widetilde{X_1}}(x)+\mu_{\widetilde{X_2}}(x))e^{\mu_{\widetilde{X_1}}(x)-\mu_{\widetilde{X_2}}(x)}-(1-\mu_{\widetilde{X_2}}(x)+\mu_{\widetilde{X_1}}(x))e^{\mu_{\widetilde{X_2}}(x)-\mu_{\widetilde{X_1}}(x)}\right).$$

In the following, we review the definition of fuzzy entropy.

Definition 2.2 [8,10]. A real function $e:\mathbf{F}(U)\to\mathbb{R}^+$ is called an entropy on $\mathbf{F}(U)$ if e possesses the following properties:

(EP1) $e(D)=0,\ \forall D\in\mathbf{P}(U)$;

(EP2) $e([\frac{1}{2}]_U)=\max_{A\in\mathbf{F}(\mathbf{U})}e(\widetilde{X_1})$;

(EP3) if $\widetilde{X_1}^*$ is a sharpened version of $\widetilde{X_1}$, then $e(\widetilde{X_1}^*)\leqslant e(X_1)$;

(EP4) $e(\widetilde{X_1}^c)=e(\widetilde{X_1}),\ \forall\widetilde{X_1}\in\mathbf{F}(U)$;

where, $[\frac{1}{2}]_U$ is the fuzzy set of U for which $\mu_{[\frac{1}{2}]_U}(x)=\frac{1}{2},\ \forall x\in U$.

Liu [24] concluded that a fuzzy entropy e is called a σ-entropy on $\mathbf{F}(U)$ if e satisfies $e(\widetilde{X})=e(\widetilde{X}\cap D)+e(\widetilde{X}\cap D^c),\ D\in\mathbf{P}(U)$.

Theorem 2.1. [24] *Let e be an entropy on* $\mathbf{F}(U)$*. Then, e is a* σ*-entropy if and only if* $\forall X_1,X_2\in\mathbf{F}(U)$*,*

$$e(\widetilde{X_1})+e(\widetilde{X_2})=e(\widetilde{X_1}\cap\widetilde{X_2})+e(\widetilde{X_1}\cup\widetilde{X_2}).$$

Fan et al. [9] introduced another theorem of σ-entropy.

Theorem 2.2. [9] *Let e be an entropy on* $\mathbf{F}(U)$*. Then e is a* σ*-entropy if and only if for* $\forall\widetilde{X}\in\mathbf{F}(U)$ *and* $D\in\mathbf{P}(U)$*,*

$$e(\widetilde{X})=e(\widetilde{X}\cap D)+e(\widetilde{X}\cap D^c).$$

In the following, common fuzzy entropies are reviewed, which are divided into three types.

(1) **Fuzzy entropies derived from the distance between a fuzzy set and its complement set**

Liu [24] presented a fuzzy entropy based on a distance as

$$e^d_{c1}(\widetilde{X})=1-d(\widetilde{X},\widetilde{X}^c),\quad\forall\widetilde{X}\in\mathbf{F}(U),$$

where d is a distance on $\mathbf{F}(U)$. If d is a σ-distance measure on $\mathbf{F}(U)$, then $e^d_{c1}(\widetilde{X})$ is a σ-entropy on $\mathbf{F}(U)$.

Fan et al. [9] defined four fuzzy entropies as

$$e^d_{c2}(\widetilde{X})=\frac{d(\widetilde{X}\cup\widetilde{X}^c,U)}{d(\widetilde{X}\cap\widetilde{X}^c,U)},\quad e^d_{c3}(\widetilde{X})=\frac{d(\widetilde{X}\cup\widetilde{X}^c,0)}{d(\widetilde{X}\cap\widetilde{X}^c,0)},\quad e^d_{c4}(\widetilde{X})=\frac{d(\widetilde{X}\cup\widetilde{X}^c,U)}{d(\widetilde{X}\cap\widetilde{X}^c,0)},\quad e^d_{c5}(\widetilde{X})=\frac{d(\widetilde{X}\cup\widetilde{X}^c,0)}{d(\widetilde{X}\cap\widetilde{X}^c,U)},$$

where the distance measure d in e^d_{c4} and e^d_{c5} should satisfy $d([1/2],U)=d([1/2],0)$. Furthermore, if d is a distance measure and satisfies $d(\widetilde{X_1},\widetilde{X_2})=d(\widetilde{X_1}^c,\widetilde{X_2}^c)$, $\widetilde{X_1},\widetilde{X_2}\in\mathbf{F}(U)$, then $e^d_{c2}(\widetilde{X})=e^d_{c3}(\widetilde{X})$ and $e^d_{c4}(\widetilde{X})=e^d_{c5}(\widetilde{X})$.

Fan and Ma [10] indicated if d is a normal distance measure, then

$$e^d_{c6}(\widetilde{X})=1-d(\widetilde{X}\cup\widetilde{X}^c,U)+d(\widetilde{X}\cap\widetilde{X}^c,U),\quad e^d_{c7}(\widetilde{X})=1-d(\widetilde{X}\cup\widetilde{X}^c,0)+d(\widetilde{X}\cap\widetilde{X}^c,0),$$

$$e^d_{c8}(\widetilde{X})=1-d(\widetilde{X}\cup\widetilde{X}^c,U)+d(\widetilde{X}\cap\widetilde{X}^c,0),\quad e^d_{c9}(\widetilde{X})=1-d(\widetilde{X}\cup\widetilde{X}^c,0)+d(\widetilde{X}\cap\widetilde{X}^c,U),$$

and if d satisfies $d(\widetilde{X_1},\widetilde{X_2})=d(\widetilde{X_1}^c,\widetilde{X_2}^c)$, $\widetilde{X_1},\widetilde{X_2}\in\mathbf{F}(U)$, then $e^d_{c6}(\widetilde{X})=e^d_{c7}(\widetilde{X})=e^d_{c8}(\widetilde{X})=e^d_{c9}(\widetilde{X})$.

Therefore, we believe that $e^d_{c2},\ e^d_{c3},\ e^d_{c4}$ and e^d_{c5} are constructed in the same manner and $e^d_{c6},\ e^d_{c7},\ e^d_{c8}$ and e^d_{c9} are similar. Therefore, we select e^d_{c2} and e^d_{c6} as the paper's examples for analysis.

(2) **Fuzzy entropy induced from a distance among a fuzzy set, its near set and its far set**

Kaufman [15] defined an entropy generated by a distance measure as

$$e^{D^p_M}_{nf1}(\widetilde{X})=\frac{2}{n^{1/p}}D^p_M(\widetilde{X},\widetilde{X}_{near}),\quad p\geqslant 1,$$

where $\widetilde{X}$ **is a fuzzy set** in $\mathbf{F}(U)$ and D^p_M is the distance measure between two fuzzy sets proposed in [17]. That $e^{D^1_M}_{nf1}$ is a σ-entropy is obvious.

Kosko [18] presented another distance-induced fuzzy entropy as

$$e_{nf2}(\widetilde{X})=\frac{D^1_M(\widetilde{X},\widetilde{X}_{near})}{D^1_M(\widetilde{X},\widetilde{X}_{far})}.$$

Fan and Xie [8] introduced a fuzzy entropy based on distance measures as

$$e^d_{nf3}(\widetilde{X}) = \frac{d(\widetilde{X}, \widetilde{X}_{near})}{d(\widetilde{X}, \widetilde{X}_{far})},$$

where d is a σ-distance and satisfies $d(\frac{1}{2}D, [0]) = d(\frac{1}{2}D, D), \ \forall D \in \mathbf{P}(U)$, $\frac{1}{2}D$ is the fuzzy set that satisfies: (1) $\mu_{\frac{1}{2}D}(x) = \frac{1}{2}$ if $x \in D$; (2) $\mu_{\frac{1}{2}D}(x) = 0$ if $x \notin D$, and $d(\widetilde{X_1}^c, \widetilde{X_2}^c) = d(\widetilde{X_1}, \widetilde{X_2}), \ \forall \widetilde{X_1}, \widetilde{X_2} \in \mathbf{F}(U)$. In fact, $e^d_{nf3}(\widetilde{X})$ is the generalization of $e_{nf2}(\widetilde{X})$.

For a set $D \subset U$,

Fan et al. [9] proposed another entropy as

$$e^d_{nf4}(\widetilde{X}) = d(\widetilde{X}, \widetilde{X}_{near}) + 1 - d(\widetilde{X}, \widetilde{X}_{far}),$$

where d is a σ-distance and satisfies $d(\frac{1}{2}D, [0]) = d(\frac{1}{2}D, D), \ \forall D \in \mathbf{P}(U)$ and $d(\widetilde{X_1}^c, \widetilde{X_2}^c) = d(\widetilde{X_1}, \widetilde{X_2}), \ \forall \widetilde{X_1}, \widetilde{X_2} \in \mathbf{F}(U)$. Fan and Ma [10] noted that $e^{D^p_M}_{nf1}(\widetilde{X}) (p \geqslant 2)$, $e_{nf2}(\widetilde{X})$ and $e^d_{nf3}(\widetilde{X})$ are generally not σ-entropies.

(3) **Other fuzzy entropies**

Fan and Ma [10] proposed a distance-induced fuzzy entropy, which is a σ-entropy.

Let d be a normal σ-distance measure on $\mathbf{F}(U)$. If d satisfies the following properties: (1) $d(\frac{1}{2}D, [0]) = d(\frac{1}{2}D, D), \ \forall D \in \mathbf{P}(U)$, (2) $d(\widetilde{X_1}^c, \widetilde{X_2}^c) = d(\widetilde{X_1}, \widetilde{X_2}), \ \forall \widetilde{X_1}, \widetilde{X_2} \in \mathbf{F}(U)$, (3) $2d(D, [\frac{1}{2}]) = 1, \ \forall D \in P(U)$, then

$$e^d_{o1}(\widetilde{X}) = 1 - 2d(\widetilde{X}, [1/2])$$

is a normal σ-entropy on $\mathbf{F}(U)$. The distance measures D_E and D_L do not satisfy the condition (3): $2d(D, [\frac{1}{2}]) = 1$.

Fan and Ma [10] proposed a σ-entropy as

$$e^{\alpha,\beta}_{o2}(\widetilde{X}) = \frac{1}{(1-\alpha)\beta} \sum_{i=1}^{n} \left((\mu_{\widetilde{X}}(x_i)^{\alpha} + (1 - \mu_{\widetilde{X}}(x_i))^{\alpha})^{\beta} - 1 \right), \quad \alpha > 0, \ \alpha \neq 1, \ \beta \neq 0.$$

Renyi [44] defined the α-order entropy of a probability distribution $(p_1, p_2, \ldots, p_n)$ as $[1/(1-\alpha)] \ln(\sum_{i=1}^{n} p^{\alpha}), \ \alpha > 0, \ \alpha \neq 1$. Similar to the above analysis, for a fuzzy set, Bhandari and Pal [3] defined the α-order entropy as

$$e^{\alpha,0}_{o2}(\widetilde{X}) = \lim_{\beta \to 0} e^{\alpha,\beta}_{6}(\widetilde{X}) = \frac{1}{1-\alpha} \sum_{i=1}^{n} \ln(\mu_{\widetilde{X}}(x_i)^{\alpha} + (1 - \mu_{\widetilde{X}}(x_i))^{\alpha}), \quad \alpha > 0, \ \alpha \neq 1.$$

Furthermore, Fan and Ma [10] provided the entropy formulas when $\beta = 1$ and $\beta = \alpha^{-1}$ as follows:

$$e^{\alpha,1}_{o2}(\widetilde{X}) = \frac{1}{1-\alpha} \sum_{i=1}^{n} \left((\mu_{\widetilde{X}}(x_i)^{\alpha} + (1 - \mu_{\widetilde{X}}(x_i))^{\alpha})^{\beta} - 1 \right), \quad \alpha > 0, \ \alpha \neq 1.$$

$$e^{\alpha,\alpha^{-1}}_{o2}(\widetilde{X}) = \frac{1}{1-\alpha} \sum_{i=1}^{n} \left((\mu_{\widetilde{X}}(x_i)^{\alpha} + (1 - \mu_{\widetilde{X}}(x_i))^{\alpha})^{\alpha^{-1}} - 1 \right), \quad \alpha > 0, \ \alpha \neq 1.$$

Shannon [45] defined the entropy of a probability distribution $(p_1, p_2, \ldots, p_n)$ as $-\sum_{i=1}^{n} p_i \ln p_i$. Applying this concept in fuzzy set theory, Luca and Termini [25] introduced a logarithmic fuzzy entropy formula for a limit universe as

$$e^k_{o3}(\widetilde{X}) = -\frac{k}{n} \sum_{i=1}^{n} \left(\mu_{\widetilde{X}}(x_i) \ln \mu_{\widetilde{X}}(x_i) + (1 - \mu_{\widetilde{X}}(x_i)) \ln(1 - \mu_{\widetilde{X}}(x_i)) \right), \quad k > 0.$$

Pal and Pal [31] analyzed the classical Shannon information entropy and introduced exponential entropy. For a probability distribution $P = (p_1, p_2, \ldots, p_n)$, exponential entropy is defined by $H = \sum_{i=1}^{n} p_i e^{1-p_i}$. For fuzzy sets, fuzzy entropy is denoted as

$$e_{o4}(\widetilde{X}) = \frac{1}{n\sqrt{e} - 1} \sum_{i=1}^{n} \left(\mu_{\widetilde{X}}(x_i) e^{1-\mu_{\widetilde{X}}(x_i)} + (1 - \mu_{\widetilde{X}}(x_i)) e^{\mu_{\widetilde{X}}(x_i)} - 1 \right).$$

Yager [48–50] proposed a complement-based fuzzy entropy, which is denoted as

$$e_{o5}(\widetilde{X}) = \frac{4}{n} \sum_{i=1}^{n} \mu_{\widetilde{X}}(x_i)(1 - \mu_{\widetilde{X}}(x_i)).$$

Based on this definition, Liang et al. [20,21] introduced the fuzzy entropy to rough set theory.

For further investigation, we regard a fuzzy entropy as a function with regard to the membership function of every object on the fuzzy set $\widetilde{X}$, i.e., $e(\widetilde{X}) = f(\mu_{\widetilde{X}}(x_1), \mu_{\widetilde{X}}(x_2), \ldots, \mu_{\widetilde{X}}(x_n))$. It is obvious that these independent variables $(\mu_{\widetilde{X}}(x_1), \mu_{\widetilde{X}}(x_2), \ldots, \mu_{\widetilde{X}}(x_n))$ are symmetric, for example $f(\mu_{\widetilde{X}}(x_1), \mu_{\widetilde{X}}(x_2), \ldots, \mu_{\widetilde{X}}(x_n)) = f(\mu_{\widetilde{X}}(x_n), \mu_{\widetilde{X}}(x_{n-1}), \ldots, \mu_{\widetilde{X}}(x_1))$.

Theorem 2.3. *Let $e(\widetilde{X}) = f(\mu_{\widetilde{X}}(x_1), \mu_{\widetilde{X}}(x_2), \ldots, \mu_{\widetilde{X}}(x_n))$. If $e(\widetilde{X})$ is a σ-entropy, then*

$$e(\widetilde{X}) = \sum_{i=1}^{n} g(\mu_{\widetilde{X}}(x_i)),$$

where $g(x_i) = f(\mu_{\widetilde{X}}(x_i), \underbrace{0, \ldots, 0}_{n-1})$.

Proof. Let $D_i = \cup_{j=1}^{i} x_j$, $\widetilde{X}_i = \sum_{j=1}^{i} \frac{0}{x_j} + \sum_{j=i+1}^{n} \frac{\mu_{\widetilde{X}}(x_j)}{x_j}$. By Theorem 2.2, $e(X) = e(\widetilde{X} \cap D) + e(\widetilde{X} \cap D^c)$ is the sufficient and necessary condition for σ-entropy. Therefore, for $\forall x_i \in U$,

$$\begin{aligned}
e(\widetilde{X}) &\iff e(\widetilde{X} \cap D_1) + e(\widetilde{X} \cap D_1^c)\\
&\iff f(\mu_{\widetilde{X}}(x_1), \underbrace{0, \ldots, 0}_{n-1}) + f(0, \mu_{\widetilde{X}}(x_2), \ldots, \mu_{\widetilde{X}}(u_n))\\
&\iff f(\mu_{\widetilde{X}}(x_1), \underbrace{0, \ldots, 0}_{n-1}) + e(\widetilde{X_1})\\
&\iff f(\mu_{\widetilde{X}}(x_1), \underbrace{0, \ldots, 0}_{n-1}) + e(\widetilde{X_1} \cap D_2) + e(\widetilde{X_1} \cap D_2^c)\\
&\iff \cdots\\
&\iff f(\mu_{\widetilde{X}}(x_1), \underbrace{0, \ldots, 0}_{n-1}) + f(0, \mu_{\widetilde{X}}(x_2), \underbrace{0, \ldots, 0}_{n-2}) + \cdots + f(\underbrace{0, \ldots, 0}_{n-1}, \mu_{\widetilde{X}}(x_n)).
\end{aligned}$$

Furthermore, we assume $g(x_i) = f(\mu_{\widetilde{X}}(x_1), \underbrace{0, \ldots, 0}_{n-1})$. Because these independent variables $(\mu_{\widetilde{X}}(x_1), \mu_{\widetilde{X}}(x_2), \ldots, \mu_{\widetilde{X}}(x_n))$ are symmetric, we have

$$\begin{aligned}
&f(\mu_{\widetilde{X}}(x_1), \underbrace{0, \ldots, 0}_{n-1}) + f(0, \mu_{\widetilde{X}}(x_2), \underbrace{0, \ldots, 0}_{n-2}) + \cdots + f(\underbrace{0, \ldots, 0}_{n-1}, \mu_{\widetilde{X}}(x_n))\\
&\iff f(\mu_{\widetilde{X}}(x_1), \underbrace{0, \ldots, 0}_{n-1}) + f(\mu_{\widetilde{X}}(x_2), \underbrace{0, \ldots, 0}_{n-1}) + \cdots + f(\mu_{\widetilde{X}}(x_n), \underbrace{0, \ldots, 0}_{n-1})\\
&\iff \sum_{i=1}^{n} g(\mu_{\widetilde{X}}(x_i)). \qquad \square
\end{aligned}$$

By Theorem 2.3, $e_{c1}^{D_M^1}$, $e_{c1}^{D_E}$, $e_{c1}^{D_L}$, $e_{c6}^{D_M^1}$, $e_{c6}^{D_E}$, $e_{c6}^{D_L}$, $e_{nf1}^{D_M^1}$, $e_{nf4}^{D_M^1}$, $e_{nf4}^{D_E}$, $e_{nf4}^{D_L}$, $e_{o1}^{D_M^1}$, $e_{o2}^{\alpha,\beta}$, e_{o3}^{k}, e_{o4} and e_{o5} are all σ-entropies.

The set-oriented view of rough sets starts from classical set algebra $(2^U, \neg, \cap, \cup)$ and associates a fuzzy set with each subset of a universe. Rough membership function may be interpreted as a special type of fuzzy membership functions, which can be interpreted in terms of probabilities defined simply by the cardinalities of sets [38,53]. In this view, a rough set Y can induce a fuzzy set $\widetilde{Y_B}$ in $\mathbf{F}(U)$ if the membership function of x_i in $\widetilde{Y_B}$ is denoted as $\mu_{\widetilde{Y_B}}(x_i) = \frac{|[x_i]_B \cap Y|}{|[x_i]_B|}$, where $Y \subseteq U$ is a crisp set and $[x_i]_B$ denotes the equivalence class of x_i determined by attribute set B. Therefore, it is possible to measure the roughness of the rough set Y by using fuzzy entropies of the fuzzy set $\widetilde{Y_B}$.

2.3. The roughness of rough sets

The roughness of a rough set results from its boundary region. For describing the properties of a rough set, Pawlak [36] discussed two numerical characterizations for evaluating a rough set's imprecision, i.e., accuracy and roughness, which are defined as follows.

Definition 2.3. [36] Let $S_A = (U, A)$ be an information system, $Y \subseteq U$. Then, the roughness of Y with respect to A is defined as

$$\rho_A(Y) = 1 - \frac{|\underline{A}(Y)|}{|\overline{A}(Y)|},$$

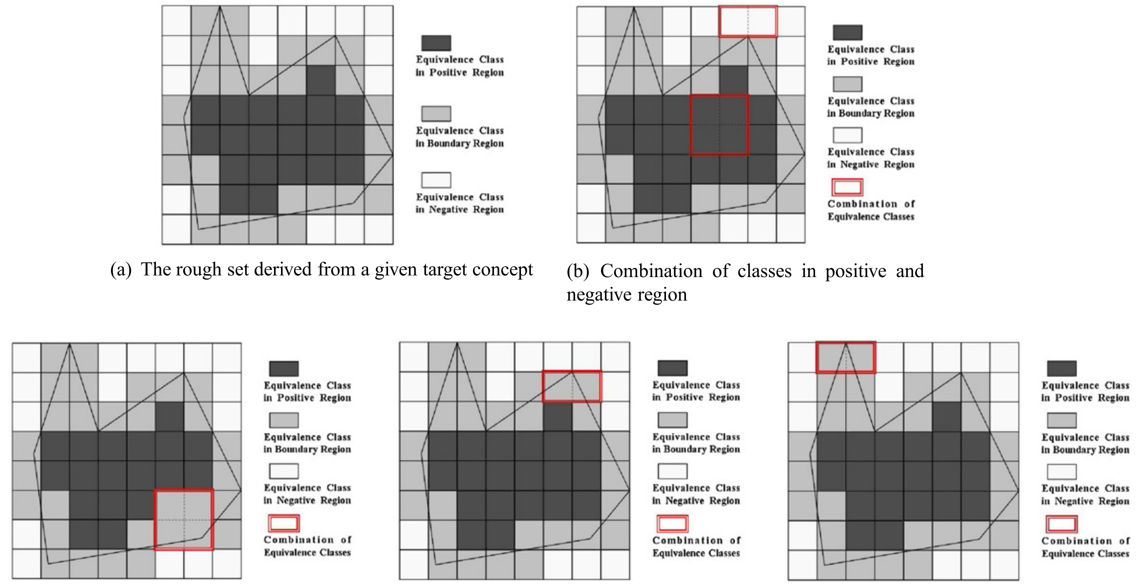
where $|\cdot|$ denotes the cardinality of a set.

By means of the definitions of rough approximations, we obtain $|\overline{A}(Y)| \leqslant |\overline{B}(Y)|$ and $|\underline{A}(Y)| \geqslant |\underline{B}(Y)|$ if $U/B \succ U/A$, i.e., $\rho_B(Y) \geqslant \rho_A(Y)$, if $U/B \succ U/A$. That is, ρ is not strictly monotonic with the partition ordering [4]. Furthermore, because $\rho_A(Y) = 1 - \frac{|\underline{A}(Y)|}{|\overline{A}(Y)|} = \frac{|\overline{A}(Y)| - |\underline{A}(Y)|}{|\overline{A}(Y)|} = \frac{|BN_A(Y)|}{|\overline{A}(Y)|}$, we obtain that the roughness measure is only determined by the size of the boundary region and upper approximation. The following example provides a concrete illustration.

Example 2.1. Let $S_1 = (U, A_1)$, $S_2 = (U, A_2)$, $S_3 = (U, A_3)$, $S_4 = (U, A_4)$, and $S_5 = (U, B)$, where $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$,
$U/A_1 = \{\{x_1, x_2, x_3, x_4, x_5, x_6\}, \{x_7, x_8, x_9, x_{10}\}, \{x_{11}, x_{12}\}\}$, $U/A_2 = \{\{x_1, x_3\}, \{x_2, x_4, x_5, x_6\}, \{x_7, x_8, x_9, x_{10}\}, \{x_{11}, x_{12}\}\}$,
$U/A_3 = \{\{x_1, x_2, x_3, x_4, x_5, x_6\}, \{x_7, x_8, x_9, x_{10}\}, \{x_{11}\}, \{x_{12}\}\}$, $U/A_4 = \{\{x_1, x_3, x_4\}, \{x_2, x_5, x_6\}, \{x_7, x_8, x_9, x_{10}\}, \{x_{11}, x_{12}\}\}$,
$U/B = \{\{x_1, x_2, x_3, x_4, x_5, x_6\}, \{x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}\}$, and $Y = \{x_1, x_2, x_7, x_{11}, x_{12}\}$.

It is clear that $U/B \succ U/A_1 \succ U/A_2$, $U/B \succ U/A_1 \succ U/A_3$ and $U/B \succ U/A_1 \succ U/A_4$. By calculation, one obtains

$$|\underline{B}(Y)| = |\emptyset| = 0, \quad |\overline{B}(Y)| = |\{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}| = 12,$$



(a) The rough set derived from a given target concept

(b) Combination of classes in positive and negative region

(c) Combination of classes in positive and boundary region

(d) Combination of classes with the different membership for the target concept in boundary region

(e) Combination of classes with the same membership for the target concept in boundary region

Fig. 1. A rough set and all of the changes derived from the combination of the equivalence classes.

$$|\underline{A_1}(Y)| = |\{x_{11}, x_{12}\}| = 2, \quad |\overline{A_1}(Y)| = |\{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}\} \cup \{x_{11}, x_{12}\}| = 12,$$

$$|\underline{A_2}(Y)| = |\{x_{11}, x_{12}\}| = 2, \quad |\overline{A_2}(Y)| = |\{x_1, x_3\} \cup \{x_2, x_4, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}\} \cup \{x_{11}, x_{12}\}| = 12,$$

$$|\underline{A_3}(Y)| = |\{x_{11}, x_{12}\}| = 2, \quad |\overline{A_3}(Y)| = |\{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}\} \cup \{x_{11}\} \cup \{x_{12}\}| = 12,$$

$$|\underline{A_4}(Y)| = |\{x_{11}, x_{12}\}| = 2, \quad |\overline{A_4}(Y)| = |\{x_1, x_3, x_4\} \cup \{x_2, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}\} \cup \{x_{11}, x_{12}\}| = 12.$$

Thus

$$|BN_B(Y)| = |\{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}| = 12,$$

$$|BN_{A_1}(Y)| = |\{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}\}| = 10,$$

$$|BN_{A_2}(Y)| = |\{x_1, x_3\} \cup \{x_2, x_4, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}\}| = 10,$$

$$|BN_{A_3}(Y)| = |\{x_1, x_2, x_3, x_4, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}\}| = 10,$$

$$|BN_{A_4}(Y)| = |\{x_1, x_3, x_4\} \cup \{x_2, x_5, x_6\} \cup \{x_7, x_8, x_9, x_{10}\}| = 10.$$

From Definition 2.3, we have that $\rho_B(Y) = \frac{|BN_A(Y)|}{|\bar{A}(Y)|} = \frac{12}{12} = 1$, $\rho_{A_1}(Y) = \frac{|BN_A(Y)|}{|\bar{A}(Y)|} = \frac{10}{12} = \frac{5}{6}$, $\rho_{A_2}(Y) = \frac{|BN_A(Y)|}{|\bar{A}(Y)|} = \frac{10}{12} = \frac{5}{6}$, $\rho_{A_3}(Y) = \frac{|BN_A(Y)|}{|\bar{A}(Y)|} = \frac{10}{12} = \frac{5}{6}$, $\rho_{A_4}(Y) = \frac{|BN_A(Y)|}{|\bar{A}(Y)|} = \frac{10}{12} = \frac{5}{6}$. Obviously, $\rho_B(Y) > \rho_{A_1}(Y) = \rho_{A_2}(Y) = \rho_{A_3}(Y) = \rho_{A_4}(Y)$.

From Example 2.1, it is clear that the structure of $BN_{A_1}(Y)$ is different from $BN_{A_2}(Y)$, even if the roughness of Y with respect to A_1 is equal to that of Y with respect to A_2. Therefore, we conclude that the roughness measure of a rough set in [36] does not contain the structural information of its boundary region.

Bianucci and Cattanco [4] obtained the same results as Example 2.1 and introduced a quantitative valuation assumed to satisfy (at least) the following two conditions:

(re1) the strict monotonicity condition: for any $Y \subseteq U$, $U/A \prec U/B$ implies $re_A(Y) < re_B(Y)$;

(re2) the boundary conditions: for any $Y \subseteq U$, $re_\omega(Y) = 0$, $re_\delta(Y) = 1$, where $\omega = \{\{x\} | x \in U\}$ and $\delta = \{U\}$.

The rough entropy proposed in [2] accords with the two conditions mentioned above, defined as follows:

Definition 2.4 [2] Let $S = (U, A)$ be an information system, $Y \subseteq U$. Then the rough entropy of Y with regard to A is defined as

$$E_A(Y) = -\rho_A(Y)\sum_{i=1}^{m}\frac{|X_i|}{n}\log_2\frac{1}{|X_i|},$$

where $n = |U|$ and $U/A = \{X_1, X_2, \ldots, X_m\}$.

Definition 2.4 shows that Beaubouef's rough entropy reflects not only the structural and size information of the boundary region of a rough set but also takes the structural and size information of its positive and negative regions into account. However, the roughness of a rough set should be evaluated by the information in its boundary region rather than in its positive and negative regions. The following example illustrates this definition.

Example 2.2 (*Continued from Example 2.1*). From the definition of Beaubouef's roughness measure, one has $E_A(Y) = -\rho_A(Y)\sum_{i=1}^{m}\frac{|X_i|}{|U|}\log_2\frac{1}{|X_i|}$. By calculation, we obtain that the values of the roughness measure $E_B(Y) = 2.5058$, $E_{A_1}(Y) = 1.7715$, $E_{A_2}(Y) = 1.3889$, $E_{A_3}(Y) = 1.6326$, $E_{A_4}(Y) = 1.3548$. Thus, $E_B(Y) > E_{A_1}(Y) > E_{A_2}(Y)$, z$E_B(Y) > E_{A_1}(Y) > E_{A_3}(Y)$ and $E_{A_1}(Y) = E_{A_4}(Y)$.

Example 2.2 indicates that the size and structure of $BN_{A_1}(Y)$ are the same as the size and structure of $BN_{A_3}(Y)$. However, the Beaubouef's rough entropy values of the rough set Y are different with respect to A_1, A_2, A_3, A_4 and B. After analysis, we find that the problem is caused by the part $\sum_{i=1}^{m}\frac{|X_i|}{n}\log_2\frac{1}{|X_i|}$, which shows that the rough entropy not only depends on the partition of the boundary region of a rough set but also on the partition of its positive and negative regions. According to the above analysis and the constraint conditions in [4], we propose a defintion of the roughness of a rough set as follows.

Definition 2.5. Let $S_1 = (U, A)$ and $S_2 = (U, B)$ be two information systems, $Y \subseteq U$. r is called as a roughness measure if r satisfies the following properties:

(RP1) $r_B(Y) = r_A(Y)$, if $BN_B(Y)/B = BN_A(Y)/A$;
(RP2) $r_B(Y) > r_A(Y)$, if $|BN_B(Y)| > |BN_A(Y)|$ and $U/B \succ U/A$;
(RP3) $r_B(Y) > r_A(Y)$, if $|BN_B(Y)| = |BN_A(Y)|$, $BN_B(Y)/B \succ BN_A(Y)/A$ and $\exists x_i \in BN_A(Y)$ such that $\mu_{\widetilde{Y_B}}(x_i) \neq \mu_{\widetilde{Y_A}}(x_i)$;
(RP4) $r_B(Y) = r_A(Y)$, if $|BN_B(Y)| = |BN_A(Y)|$, $BN_B(Y)/B \succ BN_A(Y)/A$ and $\mu_{\widetilde{Y_B}}(x_i) = \mu_{\widetilde{Y_A}}(x_i)$ for $\forall x_i \in BN_A(Y)$;
where $\mu_{\widetilde{Y_A}}(x_i) = \frac{|[x_i]_A \cap Y|}{|[x_i]_A|}$, $\mu_{\widetilde{Y_B}}(x_i) = \frac{|[x_i]_B \cap Y|}{|[x_i]_B|}$.

Fig. 1 illustrates the four properties in Definition 2.5. Fig. 1a displays the rough set derived from a partition. Fig. 1b illustrates a rough set that is approximated by a much coarser partition than in Fig. 1a. From these two figures, we can see that these two rough sets are identical, although they are under different partitions. Thus, their roughness values should be equal, as the property (RP1) shows. Fig. 1c shows a rough set that is approximated by a coarser partition than Fig. 1a. The size of its boundary region is larger than in Fig. 1a. According to these two figures, we can find that the rough set in Fig. 1c is coarser than the one in Fig. 1a. As a result, the roughness of the rough set in Fig. 1c should be greater than that in Fig. 1a, as the property (RP2) shows. Fig. 1d expresses a rough set that is approximated by a coarser partition than Fig. 1a. An equivalence class in its boundary region is generated by combining two equivalence classes whose membership degrees on the fuzzy set derived from the rough set in Fig. 1a are different. Therefore, the roughness of the rough set in Fig. 1d should be greater than that in Fig. 1a, as the property (RP3) shows. Fig. 1e displays a rough set that is approximated by a coarser partition than Fig. 1a. From it, we can see that an equivalence class in its boundary region is generated by combining two equivalence classes whose membership degrees on the fuzzy set derived from the rough set in Fig. 1a are the same. Therefore, the structure of the combined equivalence class in Fig. 1e is the same as the two shown in Fig. 1b. Therefore, we believe that the roughness of the rough set in Fig. 1e should be identical with the one in Fig. 1a, as the property (RP4) shows.

As mentioned above, roughness of a rough set arises from the existence of its boundary region. In such a boundary region, each object has a non-zero rough membership degree. When an object's rough membership degree is regarded as its fuzzy membership degree, a rough set can induce a fuzzy set. Therefore, we assert that certain fuzzy entropies may be able to evaluate the roughness of a rough set. To illustrate the inference, we use the fuzzy entropy e_{o5} in the following example.

Example 2.3 (*Continued from Example 2.1*). $\widetilde{Y_B}, \widetilde{Y_{A_1}}, \widetilde{Y_{A_2}}, \widetilde{Y_{A_3}}, \widetilde{Y_{A_4}} \in \mathbf{F}(U)$ are five fuzzy sets derived from the rough sets $(\overline{B}(Y), \underline{B}(Y))$, $(\overline{A_1}(Y), \underline{A_1}(Y))$, $(\overline{A_2}(Y), \underline{A_2}(Y))$, $(\overline{A_3}(Y), \underline{A_3}(Y))$ and $(\overline{A_4}(Y), \underline{A_4}(Y))$, respectively. They are expressed as

$$\widetilde{Y_B} = \frac{1/3}{x_1} + \frac{1/3}{x_2} + \frac{1/3}{x_3} + \frac{1/3}{x_4} + \frac{1/3}{x_5} + \frac{1/3}{x_6} + \frac{1/2}{x_7} + \frac{1/2}{x_8} + \frac{1/2}{x_9} + \frac{1/2}{x_{10}} + \frac{1/2}{x_{11}} + \frac{1/2}{x_{12}},$$

$$\widetilde{Y_{A_1}} = \frac{1/3}{x_1} + \frac{1/3}{x_2} + \frac{1/3}{x_3} + \frac{1/3}{x_4} + \frac{1/3}{x_5} + \frac{1/3}{x_6} + \frac{1/4}{x_7} + \frac{1/4}{x_8} + \frac{1/4}{x_9} + \frac{1/4}{x_{10}} + \frac{1}{x_{11}} + \frac{1}{x_{12}},$$

$$\widetilde{Y_{A_2}} = \frac{1/2}{x_1} + \frac{1/2}{x_2} + \frac{1/4}{x_3} + \frac{1/4}{x_4} + \frac{1/4}{x_5} + \frac{1/4}{x_6} + \frac{1/4}{x_7} + \frac{1/4}{x_8} + \frac{1/4}{x_9} + \frac{1/4}{x_{10}} + \frac{1}{x_{11}} + \frac{1}{x_{12}},$$

$$\widetilde{Y_{A_3}} = \frac{1/3}{x_1} + \frac{1/3}{x_2} + \frac{1/3}{x_3} + \frac{1/3}{x_4} + \frac{1/3}{x_5} + \frac{1/3}{x_6} + \frac{1/4}{x_7} + \frac{1/4}{x_8} + \frac{1/4}{x_9} + \frac{1/4}{x_{10}} + \frac{1}{x_{11}} + \frac{1}{x_{12}},$$

$$\widetilde{Y_{A_4}} = \frac{1/3}{x_1} + \frac{1/3}{x_2} + \frac{1/3}{x_3} + \frac{1/3}{x_4} + \frac{1/3}{x_5} + \frac{1/3}{x_6} + \frac{1/4}{x_7} + \frac{1/4}{x_8} + \frac{1/4}{x_9} + \frac{1/4}{x_{10}} + \frac{1}{x_{11}} + \frac{1}{x_{12}}.$$

The values of the fuzziness measure of each fuzzy set above can be calculated as follows:

$$e_{o5}(\widetilde{Y_B}) = 0.9444, \quad e_{o5}(\widetilde{Y_{A_1}}) = 0.6944, \quad e_{o5}(\widetilde{Y_{A_2}}) = 0.6667, \quad e_{o5}(\widetilde{Y_{A_3}}) = 0.6944, \quad e_{o5}(\widetilde{Y_{A_4}}) = 0.6944.$$

Obviously, we have

$$e_{o5}(\widetilde{Y_{A_1}}) = e_{o5}(\widetilde{Y_{A_3}}),$$

when $BN_{A_1}(Y)/A_1 = BN_{A_3}(Y)/A_3$ & $|BN_{A_1}(Y)/A_1| = |BN_{A_3}(Y)/A_3|$, which reflects that e_{o5} accords with (RP1);

$$e_{o5}(\widetilde{Y_B}) > e_{o5}(\widetilde{Y_{A_1}}), \quad e_{o5}(\widetilde{Y_B}) > e_{o5}(\widetilde{Y_{A_2}}), \quad e_{o5}(\widetilde{Y_B}) > e_{o5}(\widetilde{Y_{A_3}}) \quad \& \; e_{o5}(\widetilde{Y_B}) > e_{o5}(\widetilde{Y_{A_4}}),$$

when $U/B \succ U/A_1$ & $|BN_B(Y)/B| > |BN_{A_1}(Y)/A_1|$, $U/B \succ U/A_2$ & $|BN_B(Y)/B| > |BN_{A_2}(Y)/A_2|$, $U/B \succ U/A_3$ & $|BN_B(Y)/B| > |BN_{A_3}(Y)/A_3|$, $U/B \succ U/A_4$ & $|BN_B(Y)/B| > |BN_{A_4}(Y)/A_4|$, which illustrates that e_{o5} is coincident with (RP2);

$$e_{o5}(\widetilde{Y_{A_1}}) > e_{o5}(\widetilde{Y_{A_2}}),$$

when $BN_{A_1}(Y)/A_1 \succ BN_{A_2}(Y)/A_2$ & $|BN_{A_1}(Y)/A_1| = |BN_{A_2}(Y)/A_2|$ & $\mu_{\widetilde{Y_{A_1}}}(x_1) = \mu_{\widetilde{Y_{A_1}}}(x_2) \neq \mu_{\widetilde{Y_{A_2}}}(x_1) = \mu_{\widetilde{Y_{A_2}}}(x_2)$ & $\mu_{\widetilde{Y_{A_1}}}(x_3) = \mu_{\widetilde{Y_{A_1}}}(x_4) = \mu_{\widetilde{Y_{A_1}}}(x_5) = \mu_{\widetilde{(Y_{A_1}}}(x_6) \neq \mu_{\widetilde{Y_{A_2}}}(x_3) = \mu_{\widetilde{Y_{A_2}}}(x_4) = \mu_{\widetilde{Y_{A_2}}}(x_5) = \mu_{\widetilde{Y_{A_2}}}(x_6)$, which shows that e_{o5} accords with (RP3);

$$e_{o5}(\widetilde{Y_{A_1}}) = e_{o5}(\widetilde{Y_{A_4}}),$$

when $BN_{A_1}(Y)/A_1 \succ BN_{A_4}(Y)/A_4$ & $|BN_{A_1}(Y)/A_1| = |BN_{A_4}(Y)/A_4|$ & $\mu_{\widetilde{Y_{A_1}}}(x_1) = \cdots = \mu_{\widetilde{Y_{A_1}}}(x_{12}) = \mu_{\widetilde{Y_{A_4}}}(x_1) = \cdots = \mu_{\widetilde{Y_{A_4}}}(x_{12})$, which shows that e_{o5} accords with (RP4).

The results in Example 2.3 confirm that our estimation is correct. Thus, in the remainder of the paper, we examine further the question "Can fuzzy entropies be used to evaluate the roughness of a rough set?"

3. Methods of judging whether fuzzy entropies evaluate roughness

In this section, we first investigate the change mechanism of fuzzy entropies varying with partition ordering. Next, based on the mechanism, we propose three methods of determining whether fuzzy entropies can be used to evaluate the roughness of a rough set.

For convenience, in the following, a fuzzy entropy $e(\widetilde{X})$ is regarded as the function determined by n independent variables $(\mu_{\widetilde{X}}(x_1), \mu_{\widetilde{X}}(x_2), \ldots, \mu_{\widetilde{X}}(x_n))$. Thus, it is straightforward to formulate a fuzzy entropy as $e(\widetilde{X}) = f(\mu_{\widetilde{X}}(x_1), \mu_{\widetilde{X}}(x_2), \ldots, \mu_{\widetilde{X}}(x_n))$, where $\widetilde{X}$ is a fuzzy set on $\mathbf{F}(U)$.

Definition 3.1. Given a fuzzy set on $\mathbf{F}(U)$, we denote its fold set as

$$\mu_{\widetilde{X}^F}(x) = \begin{cases} \mu_{\widetilde{X}}(x), & \mu_{\widetilde{X}}(x) \leqslant \frac{1}{2}, \\ 1 - \mu_{\widetilde{X}}(x), & \mu_{\widetilde{X}}(x) > \frac{1}{2}, \end{cases}$$

where $\mu_{\widetilde{X}}(x)$ is the membership function of x in $\widetilde{X} \in \mathbf{F}(U)$, $x \in U$.

In the following, we will analyze the relationship between $e(\widetilde{X})$ and $e(\widetilde{X}^F)$.

Property 3.1. *Let $\widetilde{X} \in \mathbf{F}(U)$ be an arbitrary fuzzy set in U. If the fuzzy entropy e is a σ-entropy, then*

$$e(\widetilde{X}) = e(\widetilde{X}^F).$$

Proof. Without any loss of generality, we suppose that for one fuzzy set $\widetilde{X}$, $\mu_{\widetilde{X}}(x_i) \leqslant \frac{1}{2}(1 \leqslant i \leqslant p)$ and $\mu_{\widetilde{X}}(x_j) > \frac{1}{2}(p < j \leqslant n)$. Then $\mu_{\widetilde{X}^F}(x_i) = \mu_{\widetilde{X}}(x_i)$, $\mu_{\widetilde{X}^F_{near}}(x_i) = \mu_{\widetilde{X}_{near}}(x_i)$, $\mu_{\widetilde{X}^F}(x_j) = 1 - \mu_{\widetilde{X}}(x_j)$, $\mu_{\widetilde{X}^F_{near}}(x_j) = 1 - \mu_{\widetilde{X}_{near}}(x_j)$. Let $e_\sigma(\widetilde{X}^F) = f(\mu_{\widetilde{X}^F}(x_1), \ldots, \mu_{\widetilde{X}^F}(x_n))$. By Theorem 2.2, we have

$$\begin{aligned} e(\widetilde{X}^F) &= e(\widetilde{X}^F \cap D) + e(\widetilde{X}^F \cap D^c) \\ &= f(\mu_{\widetilde{X}^F}(x_1), \ldots, \mu_{\widetilde{X}^F}(x_p), \underbrace{0, \ldots, 0}_{n-p}) + f(\underbrace{0, \ldots, 0}_{p}, \mu_{\widetilde{X}^F}(x_{p+1}), \ldots, \mu_{\widetilde{X}^F}(x_n)) \\ &= f(\mu_{\widetilde{X}}(x_1), \ldots, \mu_{\widetilde{X}}(x_p), \underbrace{0, \ldots, 0}_{n-p}) + f(\underbrace{0, \ldots, 0}_{p}, 1 - \mu_{\widetilde{X}}(x_{p+1}), \ldots, 1 - \mu_{\widetilde{X}}(x_n)). \end{aligned}$$

Furthermore, let $D_1 = \{x_1, x_2, \ldots, x_p\}$, $\widetilde{F_1} = \frac{\mu_{\widetilde{X}}(x_1)}{x_1} + \cdots + \frac{\mu_{\widetilde{X}}(x_p)}{x_p} + \frac{0}{x_{p+1}} + \cdots + \frac{0}{x_n}$ and $\widetilde{F_2} = \frac{0}{x_1} + \cdots + \frac{0}{x_p} + \frac{1-\mu_{\widetilde{X}}(x_{p+1})}{x_{p+1}} + \cdots + \frac{1-\mu_{\widetilde{X}}(x_n)}{x_n}$. Then $\widetilde{F_2}^c = \frac{1}{x_1} + \cdots + \frac{1}{x_p} + \frac{\mu_{\widetilde{X}}(x_{p+1})}{x_{p+1}} + \cdots + \frac{\mu_{\widetilde{X}}(x_n)}{x_n}$. From Theorem 2.1, we have

$$e(\widetilde{F_1}) + e(\widetilde{F_2}) = e(\widetilde{F_1}) + e(\widetilde{F_2}^c) = e(\widetilde{F_1} \cap \widetilde{F_2}^c) + e(\widetilde{F_1} \cup \widetilde{F_2}^c) = e(D_1) + e(\widetilde{X}) = e(\widetilde{X}).$$

Thus, $e(\widetilde{X}^F) = e(\widetilde{X})$. □

For a non-σ-entropy, we provide the following properties to indicate the relationship between $e(\widetilde{X})$ and $e(\widetilde{X}^F)$.

Property 3.2. *Let $\widetilde{X} \in \mathbf{F}(U)$ be an arbitrary fuzzy set in U. Then*

$$e^d_{c2}(\widetilde{X}) = e^d_{c2}(\widetilde{X}^F).$$

Proof. Without any loss of generality, we suppose that for one fuzzy set $\widetilde{X} \in \mathbf{F}(U)$, $\mu_{\widetilde{X}}(x_i) \leqslant \frac{1}{2}(1 \leqslant i \leqslant p)$ and $\mu_{\widetilde{X}}(x_j) > \frac{1}{2}(p < j \leqslant n)$. Then,

$$\begin{aligned}
&\widetilde{X} = \mu_{\widetilde{X}}(x_1)/x_1 + \cdots + \mu_{\widetilde{X}}(x_p)/x_p + \mu_{\widetilde{X}}(x_{p+1})/x_{p+1} + \cdots + \mu_{\widetilde{X}}(x_n)/x_n,\\
&\widetilde{X}^c = (1-\mu_{\widetilde{X}}(x_1))/x_1 + \cdots + (1-\mu_{\widetilde{X}}(x_p))/x_p + (1-\mu_{\widetilde{X}}(x_{p+1}))/x_{p+1} + \cdots + (1-\mu_{\widetilde{X}}(x_n))/x_n,\\
&\widetilde{X}^F = \mu_{\widetilde{X}}(x_1)/x_1 + \cdots + \mu_{\widetilde{X}}(x_p)/x_p + (1-\mu_{\widetilde{X}}(x_{p+1}))/x_{p+1} + \cdots + (1-\mu_{\widetilde{X}}(x_n))/x_n,\\
&(\widetilde{X}^F)^c = (1-\mu_{\widetilde{X}}(x_1))/x_1 + \cdots + (1-\mu_{\widetilde{X}}(x_p))/x_p + \mu_{\widetilde{X}}(x_{p+1})/x_{p+1} + \cdots + \mu_{\widetilde{X}}(x_n)/x_n.
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
&\widetilde{X} \cap \widetilde{X}^c = \mu_{\widetilde{X}}(x_1)/x_1 + \cdots + \mu_{\widetilde{X}}(x_p)/x_p + (1-\mu_{\widetilde{X}}(x_{p+1}))/x_{p+1} + \cdots + (1-\mu_{\widetilde{X}}(x_n))/x_n,\\
&X^F \cap (\widetilde{X}^F)^c = \mu_{\widetilde{X}}(x_1)/x_1 + \cdots + \mu_{\widetilde{X}}(x_p)/x_p + (1-\mu_{\widetilde{X}}(x_{p+1}))/x_{p+1} + \cdots + (1-\mu_{\widetilde{X}}(x_n))/x_n,\\
&\widetilde{X} \cup \widetilde{X}^c = (1-\mu_{\widetilde{X}}(x_1))/x_1 + \cdots + (1-\mu_{\widetilde{X}}(x_p))/x_p + \mu_{\widetilde{X}}(x_{p+1})/x_{p+1} + \cdots + \mu_{\widetilde{X}}(x_n)/x_n,\\
&\widetilde{X}^F \cup (\widetilde{X}^F)^c = (1-\mu_{\widetilde{X}}(x_1))/x_1 + \cdots + (1-\mu_{\widetilde{X}}(x_p))/x_p + \mu_{\widetilde{X}}(x_{p+1})/x_{p+1} + \cdots + \mu_{\widetilde{X}}(x_n)/x_n.
\end{aligned}$$

It is obvious that $\widetilde{X} \cap \widetilde{X}^c = \widetilde{X}^F \cap (X^F)^c$ and $\widetilde{X} \cup \widetilde{X}^c = \widetilde{X}^F \cup (\widetilde{X}^F)^c$. Therefore,

$$e^d_{c2}(\widetilde{X}^F) = \frac{d(\widetilde{X}^F \cup (\widetilde{X}^F)^c, U)}{d(\widetilde{X}^F \cap (\widetilde{X}^F)^c, U)} = \frac{d(\widetilde{X} \cup \widetilde{X}^c, U)}{d(\widetilde{X} \cap \widetilde{X}^c, U)} = e^d_{c2}(\widetilde{X}). \qquad \square$$

Property 3.2 states that $e^{D^p_M}_{c2}(\widetilde{X}^F)$, $e^{D_E}_{c2}(\widetilde{X}^F)$ and $e^{D_L}_{c2}(\widetilde{X}^F)$ are the same as $e^{D^p_M}_{c2}(\widetilde{X}^F)$, $e^{D_E}_{c2}(\widetilde{X}^F)$ and $e^{D_L}_{c2}(\widetilde{X}^F)$, respectively.

Property 3.3. *Let $\widetilde{X} \in \mathbf{F}(U)$ be an arbitrary fuzzy set. Then,*

$$e^{D^p_M}_{nf1}(\widetilde{X}) = e^{D^p_M}_{nf1}(\widetilde{X}^F), \quad e^{D^p_M}_{nf3}(\widetilde{X}) = e^{D^p_M}_{nf3}(\widetilde{X}^F), \quad e^{D_E}_{nf3}(\widetilde{X}) = e^{D_E}_{nf3}(\widetilde{X}^F), \quad e^{D_L}_{nf3}(\widetilde{X}) = e^{D_L}_{nf3}(\widetilde{X}^F).$$

Proof. Without any loss of generality, we suppose that for one fuzzy set $\widetilde{X}$, $\mu_{\widetilde{X}}(x_i) \leqslant \frac{1}{2}(1 \leqslant i \leqslant p)$ and $\mu_{\widetilde{X}}(x_j) > \frac{1}{2}(p < j \leqslant n)$. Then $\mu_{\widetilde{X}^F}(x_i) = \mu_{\widetilde{X}}(x_i)$, $\mu_{\widetilde{X}^F_{near}}(x_i) = \mu_{\widetilde{X}_{near}}(x_i) = 0$, $\mu_{\widetilde{X}^F_{far}}(x_i) = \mu_{\widetilde{X}_{far}}(x_i) = 1$, $\mu_{\widetilde{X}^F}(x_j) = 1 - \mu_{\widetilde{X}}(u_j)$, $\mu_{\widetilde{X}^F_{near}}(x_j) = 1 - \mu_{\widetilde{X}_{near}}(x_j) = 1$, $\mu_{\widetilde{X}^F_{far}}(x_j) = 1 - \mu_{\widetilde{X}_{far}}(x_j) = 0$.

Therefore, we obtain

$$\begin{aligned}
e^p_{nf1}(\widetilde{X}^F) &= \frac{2}{n^{1/p}}\left(\sum_{i=1}^{n}|\mu_{\widetilde{X}^F}(x_i) - \mu_{\widetilde{X}^F_{near}}(x_i)|^p\right)^{1/p}\\
&= \frac{2}{n^{1/p}}\left(\sum_{i=1}^{p}|\mu_{\widetilde{X}}(x_i) - \mu_{\widetilde{X}_{near}}(x_i)|^p + \sum_{j=p+1}^{n}|1 - \mu_{\widetilde{X}}(x_j) - \left(1 - \mu_{\widetilde{X}_{near}}(x_j)\right)|^p\right)^{1/p}\\
&= \frac{2}{n^{1/p}}\left(\sum_{i=1}^{n}|\mu_{\widetilde{X}}(x_i) - \mu_{\widetilde{X}_{near}}(x_i)|^p\right)^{1/p}\\
&= e^p_{nf1}(\widetilde{X})(p \geqslant 1).
\end{aligned}$$

Similarly, $e^{D^p_M}_{nf3}(\widetilde{X}) = e^{D^p_M}_{nf3}(\widetilde{X}^F)$, $e^{D_E}_{nf3}(\widetilde{X}) = e^{D_E}_{nf3}(\widetilde{X}^F)$, $e^{D_L}_{nf3}(\widetilde{X}) = e^{D_L}_{nf3}(\widetilde{X}^F)$ can be proven. □

Table 2
Roughness measures of the decision class in the dataset Spect.

Measure	Class	Number of condition attribute										
		22	20	18	16	14	12	10	8	6	4	2
ρ	1	0.2033	0.2358	0.2619	0.2698	0.3110	0.3789	0.4154	0.4771	0.6391	0.7903	1.0000
	2	0.6533	0.7342	0.8148	0.8193	0.8587	0.8981	0.9391	0.9615	0.9942	1.0000	1.0000
E	1	0.0165	0.0258	0.0345	0.0392	0.0540	0.0814	0.1048	0.1503	0.2841	0.4503	0.7812
	2	0.0530	0.0805	0.1074	0.1190	0.1490	0.1930	0.2370	0.3030	0.4419	0.5698	0.7812
$e_{c1}^{D_M^1}$	1	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
	2	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
$e_{c1}^{D_M^2}$	1	0.0802	0.1054	0.1215	0.1264	0.1452	0.1673	0.1865	0.2155	0.3006	0.3292	0.3741
	2	0.0802	0.1054	0.1215	0.1264	0.1452	0.1673	0.1865	0.2155	0.3006	0.3292	0.3741
$e_{c1}^{D_E}$	1	0.1586	0.2030	0.2317	0.2401	0.2741	0.3146	0.3465	0.3957	0.5267	0.5740	0.6445
	2	0.1586	0.2030	0.2317	0.2401	0.2741	0.3146	0.3465	0.3957	0.5267	0.5740	0.6445
$e_{c1}^{D_L}$	1	0.1548	0.2003	0.2288	0.2374	0.2702	0.3081	0.3398	0.3865	0.5137	0.5545	0.6149
	2	0.1548	0.2003	0.2288	0.2374	0.2702	0.3081	0.3398	0.3865	0.5137	0.5545	0.6149
$e_{c2}^{D_M^1}$	1	0.0637	0.0898	0.1033	0.1079	0.1218	0.1362	0.1558	0.1711	0.2535	0.2594	0.2594
	2	0.0637	0.0898	0.1033	0.1079	0.1218	0.1362	0.1558	0.1711	0.2535	0.2594	0.2594
$e_{c2}^{D_M^2}$	1	0.1542	0.1934	0.2076	0.2128	0.2238	0.2319	0.2538	0.2567	0.3335	0.3230	0.2898
	2	0.1542	0.1934	0.2076	0.2128	0.2238	0.2319	0.2538	0.2567	0.3335	0.3230	0.2898
$e_{c2}^{D_E}$	1	0.0198	0.0313	0.0360	0.0379	0.0420	0.0452	0.0543	0.0556	0.0960	0.0904	0.0728
	2	0.0198	0.0313	0.0360	0.0379	0.0420	0.0452	0.0543	0.0556	0.0960	0.0904	0.0728
$e_{c2}^{D_L}$	1	0.0241	0.0378	0.0435	0.0457	0.0507	0.0545	0.0652	0.0670	0.1128	0.1062	0.0866
	2	0.0241	0.0378	0.0435	0.0457	0.0507	0.0545	0.0652	0.0670	0.1128	0.1062	0.0866
$e_{c6}^{D_M^1}$	1	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
	2	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
$e_{c6}^{D_M^2}$	1	0.1969	0.2487	0.2697	0.2770	0.2954	0.3115	0.3407	0.3530	0.4545	0.4514	0.4310
	2	0.1969	0.2487	0.2697	0.2770	0.2954	0.3115	0.3407	0.3530	0.4545	0.4514	0.4310
$e_{c6}^{D_E}$	1	0.1252	0.1698	0.1931	0.2006	0.2247	0.2498	0.2799	0.3061	0.4212	0.4346	0.4448
	2	0.1252	0.1698	0.1931	0.2006	0.2247	0.2498	0.2799	0.3061	0.4212	0.4346	0.4448
$e_{c6}^{D_L}$	1	0.1208	0.1657	0.1883	0.1958	0.2186	0.2415	0.2715	0.2946	0.4075	0.4160	0.4179
	2	0.1208	0.1657	0.1883	0.1958	0.2186	0.2415	0.2715	0.2946	0.4075	0.4160	0.4179
$e_{nf1}^{D_M^1}$	1	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
	2	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
$e_{nf1}^{D_M^2}$	1	0.2929	0.3603	0.3826	0.3908	0.4063	0.4157	0.4484	0.4470	0.5461	0.5234	0.4644
	2	0.2929	0.3603	0.3826	0.3908	0.4063	0.4157	0.4484	0.4470	0.5461	0.5234	0.4644
e_{nf2}	1	0.0637	0.0898	0.1033	0.1079	0.1218	0.1362	0.1558	0.1711	0.2535	0.2594	0.2594
	2	0.0637	0.0898	0.1033	0.1079	0.1218	0.1362	0.1558	0.1711	0.2535	0.2594	0.2594
$e_{nf3}^{D_E}$	1	0.0198	0.0313	0.0360	0.0379	0.0420	0.0452	0.0543	0.0556	0.096	0.0904	0.0728
	2	0.0198	0.0313	0.0360	0.0379	0.0420	0.0452	0.0543	0.0556	0.096	0.0904	0.0728
$e_{nf3}^{D_L}$	1	0.0241	0.0378	0.0435	0.0457	0.0507	0.0545	0.0652	0.0670	0.1128	0.1062	0.0866
	2	0.0241	0.0378	0.0435	0.0457	0.0507	0.0545	0.0652	0.0670	0.1128	0.1062	0.0866
$e_{nf4}^{D_M^1}$	1	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
	2	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
$e_{nf4}^{D_E}$	1	0.1252	0.1698	0.1931	0.2006	0.2247	0.2498	0.2799	0.3061	0.4212	0.4346	0.4448
	2	0.1252	0.1698	0.1931	0.2006	0.2247	0.2498	0.2799	0.3061	0.4212	0.4346	0.4448
$e_{nf1}^{D_L}$	1	0.1252	0.1698	0.1931	0.2006	0.2247	0.2498	0.2799	0.3061	0.4212	0.4346	0.4448
	2	0.1252	0.1698	0.1931	0.2006	0.2247	0.2498	0.2799	0.3061	0.4212	0.4346	0.4448
e_{o1}	1	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
	2	0.1199	0.1648	0.1873	0.1948	0.2172	0.2397	0.2697	0.2921	0.4045	0.4120	0.4120
$e_{o2}^{\alpha,\beta}$	1	0.1406	0.1736	0.1978	0.2044	0.2352	0.2781	0.3071	0.3532	0.4752	0.5469	0.6473
	2	0.1406	0.1736	0.1978	0.2044	0.2352	0.2781	0.3071	0.3532	0.4752	0.5469	0.6473
	1	0.1117	0.1417	0.1616	0.1674	0.1915	0.2215	0.2441	0.2793	0.3730	0.4126	0.4699

(continued on next page)

Table 2 (*continued*)

Measure	Class	Number of condition attribute										
		22	20	18	16	14	12	10	8	6	4	2
e_{o3}^k												
	2	0.1117	0.1417	0.1616	0.1674	0.1915	0.2215	0.2441	0.2793	0.3730	0.4126	0.4699
e_{o4}	1	0.0611	0.0791	0.0903	0.0937	0.1066	0.1216	0.1341	0.1526	0.2028	0.2190	0.2430
	2	0.0611	0.0791	0.0903	0.0937	0.1066	0.1216	0.1341	0.1526	0.2028	0.2190	0.2430
e_{o5}	1	0.1539	0.1998	0.2282	0.2368	0.2694	0.3066	0.3383	0.3845	0.5108	0.5500	0.6083
	2	0.1539	0.1998	0.2282	0.2368	0.2694	0.3066	0.3383	0.3845	0.5108	0.5500	0.6083

Based on the above properties, in the remainder of the paper, we change the focus of the investigation from the change mechanism $e(\widetilde{X})$ to the change mechanism $e(\widetilde{X}^F)$.

To provide a method of determining whether fuzzy entropies evaluate roughness, we introduce the following four theorems.

Theorem 3.1. *Let $S_1 = (U,A)$ and $S_2 = (U,B)$ be two information systems, and let $Y \subseteq U$ and $\widetilde{Y_A}, \widetilde{Y_B} \in \mathbf{F}(U)$ be two fuzzy sets. If $[x_p]_B = [x_p]_A \cup [x_q]_A$ ($x_p, x_q \in U$), for $\forall x_i \notin [x_p]_B$ such that $[x_i]_B = [x_i]_A$, and $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q) \neq \mu_{\widetilde{Y_B}}(x_p)$, then*

$$|[u_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) \leqslant 0.$$

In particular, if $\mu_{\widetilde{Y_B}}(x_p) \leqslant \frac{1}{2}$ & $\mu_{\widetilde{Y_A}}(x_p) \leqslant \frac{1}{2}$ & $\mu_{\widetilde{Y_A}}(x_q) \leqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_B}}(x_p) \geqslant \frac{1}{2}$ & $\mu_{\widetilde{Y_A}}(x_p) \geqslant \frac{1}{2}$ & $\mu_{\widetilde{Y_A}}(x_q) \geqslant \frac{1}{2}$,

$$|[x_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(u_q) = 0,$$

where $\mu_{\widetilde{Y_A}}(x_i) = \frac{|[x_i]_A \cap Y|}{|[x_i]_A|}$, $\mu_{\widetilde{Y_B}}(x_i) = \frac{|[x_i]_B \cap Y|}{|[x_i]_B|}$, and $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i) = \mu_{\widetilde{Y_A}^F}(x_i) - \mu_{\widetilde{Y_B}^F}(x_i)$.

Proof. To prove the theorem, six cases should be investigated as follows.

(1) $\mu_{\widetilde{Y_B}}(x_p) \leqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \leqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \leqslant \frac{1}{2}$

$$\begin{aligned}\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) &= \mu_{\widetilde{Y_A}^F}(x_p) - \mu_{\widetilde{Y_B}^F}(x_p)\\ &= \mu_{\widetilde{Y_A}}(x_p) - \mu_{\widetilde{Y_B}}(x_p)\\ &= \frac{|[x_p]_A \cap Y|}{|[x_p]_A|} - \frac{|[x_p]_B \cap Y|}{|[x_p]_B|}\\ &= \frac{|[x_p]_A \cap Y|}{|[x_p]_A|} - \frac{|([x_p]_A \cup [x_q]_A) \cap Y|}{|[x_p]_A \cup [u_q]_A|}\\ &= \frac{|[x_p]_A \cap Y| \times |[x_p]_A \cup [x_q]_A|}{|[x_p]_A \cup [x_q]_A| \times |[x_p]_A|} - \frac{|[x_p]_A| \times |([x_p]_A \cup [x_q]_A) \cap Y|}{|[x_p]_A \cup [x_q]_A| \times |[x_p]_A|}\\ &= \frac{|[x_p]_A \cap Y| \times |[x_q]_A| - |[x_p]_A| \times |[x_q]_A \cap Y|}{(|[x_p]_A| + |[x_q]_A|) \times |[x_p]_A|}.\end{aligned}$$

Similarly, we have

$$\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = \frac{|[x_q]_A \cap Y| \times |[x_p]_A| - |[x_q]_A| \times |[x_p]_A \cap Y|}{(|[x_q]_A| + |[x_p]_A|) \times |[x_q]_A|}.$$

Thus,

$$|[u_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = 0.$$

(2) $\mu_{\widetilde{Y_B}}(x_p) \geqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \geqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \geqslant \frac{1}{2}$

$$\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) = \mu_{\widetilde{Y_A}^F}(x_p) - \mu_{\widetilde{Y_B}^F}(x_p) = 1 - \mu_{\widetilde{Y_A}}(x_p) - (1 - \mu_{\widetilde{Y_B}}(x_p)) = \mu_{\widetilde{Y_B}}(x_p) - \mu_{\widetilde{Y_A}}(x_p).$$

Similarly,

$$\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = \mu_{\widetilde{Y_B}}(x_q) - \mu_{\widetilde{Y_A}}(x_p).$$

According to Case (1),

$$|[x_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(_q) = 0.$$

Table 3
Roughness measures of the decision class in the datasetKr-vs-kp

Measure	Class	Number of condition attribute								
		36	32	28	24	20	16	12	8	4
ρ	1	0.0000	0.3738	0.4052	0.4546	0.6907	0.7620	0.9226	0.9816	0.9991
	2	0.0000	0.3850	0.4173	0.4711	0.6993	0.7629	0.8897	0.9535	0.9997
E	1	0.0000	0.0302	0.0382	0.0639	0.1579	0.2292	0.3905	0.6018	0.8660
	2	0.0000	0.0311	0.0393	0.0663	0.1599	0.2294	0.3766	0.5846	0.8666
$e_{c1}^{D_M^1}$	1	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
	2	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
$e_{c1}^{D_M^2}$	1	0.0000	0.1148	0.1249	0.1507	0.2513	0.2961	0.4178	0.5819	0.8519
	2	0.0000	0.1148	0.1249	0.1507	0.2513	0.2961	0.4178	0.5819	0.8519
$e_{c1}^{D_E}$	1	0.0000	0.2195	0.2380	0.2826	0.4522	0.5189	0.6795	0.8393	0.9814
	2	0.0000	0.2195	0.2380	0.2826	0.4522	0.5189	0.6795	0.8393	0.9814
$e_{c1}^{D_L}$	1	0.0000	0.2170	0.2349	0.2794	0.4418	0.5072	0.6644	0.8277	0.9787
	2	0.0000	0.2170	0.2349	0.2794	0.4418	0.5072	0.6644	0.8277	0.9787
$e_{c2}^{D_M^1}$	1	0.0000	0.1036	0.1117	0.1337	0.2111	0.2509	0.3588	0.5212	0.8336
	2	0.0000	0.1036	0.1117	0.1337	0.2111	0.2509	0.3588	0.5212	0.8336
$e_{c2}^{D_M^2}$	1	0.0000	0.2160	0.2232	0.2438	0.3000	0.3325	0.4161	0.5497	0.8357
	2	0.0000	0.2160	0.2232	0.2438	0.3000	0.3325	0.4161	0.5497	0.8357
$e_{c2}^{D_E}$	1	0.0000	0.0393	0.0420	0.0501	0.0770	0.0953	0.153	0.2765	0.6814
	2	0.0000	0.0393	0.0420	0.0501	0.0770	0.0953	0.153	0.2765	0.6814
$e_{c2}^{D_L}$	1	0.0000	0.0470	0.0502	0.0599	0.0912	0.1121	0.1757	0.3064	0.7026
	2	0.0000	0.0470	0.0502	0.0599	0.0912	0.1121	0.1757	0.3064	0.7026
$e_{c6}^{D_M^1}$	1	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
	2	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
$e_{c6}^{D_M^2}$	1	0.0000	0.2763	0.2876	0.3184	0.4077	0.4523	0.5589	0.6976	0.9099
	2	0.0000	0.2763	0.2876	0.3184	0.4077	0.4523	0.5589	0.6976	0.9099
$e_{c6}^{D_E}$	1	0.0000	0.1920	0.2059	0.2420	0.3627	0.4172	0.5487	0.7054	0.9180
	2	0.0000	0.1920	0.2059	0.2420	0.3627	0.4172	0.5487	0.7054	0.9180
$e_{c6}^{D_L}$	1	0.0000	0.1885	0.2018	0.2370	0.3511	0.4040	0.5318	0.6888	0.9108
	2	0.0000	0.1885	0.2018	0.2370	0.3511	0.4040	0.5318	0.6888	0.9108
$e_{nf1}^{D_M^1}$	1	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
	2	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
$e_{nf1}^{D_M^2}$	1	0.0000	0.3987	0.4093	0.4395	0.5076	0.5457	0.6287	0.7384	0.9168
	2	0.0000	0.3987	0.4093	0.4395	0.5076	0.5457	0.6287	0.7384	0.9168
e_{nf2}	1	0.0000	0.1036	0.1117	0.1337	0.2111	0.2509	0.3588	0.5212	0.8336
	2	0.0000	0.1036	0.1117	0.1337	0.2111	0.2509	0.3588	0.5212	0.8336
$e_{nf3}^{D_E}$	1	0.0000	0.0393	0.0420	0.0501	0.0770	0.0953	0.1530	0.2765	0.6814
	2	0.0000	0.0393	0.0420	0.0501	0.0770	0.0953	0.1530	0.2765	0.6814
$e_{nf3}^{D_L}$	1	0.0000	0.0470	0.0502	0.0599	0.0912	0.1121	0.1757	0.3064	0.7026
	2	0.0000	0.0470	0.0502	0.0599	0.0912	0.1121	0.1757	0.3064	0.7026
$e_{nf4}^{D_M^1}$	1	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
	2	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
$e_{nf4}^{D_E}$	1	0.0000	0.1920	0.2059	0.2420	0.3627	0.4172	0.5487	0.7054	0.9180
	2	0.0000	0.1920	0.2059	0.2420	0.3627	0.4172	0.5487	0.7054	0.9180
$e_{nf1}^{D_L}$	1	0.0000	0.1920	0.2059	0.2420	0.3627	0.4172	0.5487	0.7054	0.9180
	2	0.0000	0.1920	0.2059	0.2420	0.3627	0.4172	0.5487	0.7054	0.9180
e_{o1}	1	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
	2	0.0000	0.1877	0.2009	0.2359	0.3486	0.4011	0.5282	0.6852	0.9093
$e_{o2}^{\alpha,\beta}$	1	0.0000	0.1873	0.2047	0.2410	0.4033	0.4639	0.6126	0.7286	0.8194
	2	0.0000	0.1873	0.2047	0.2410	0.4033	0.4639	0.6126	0.7286	0.8194
	1	0.0000	0.1531	0.1664	0.1971	0.3191	0.3665	0.4810	0.5882	0.6815

(*continued on next page*)

Table 3 (*continued*)

Measure	Class	Number of condition attribute								
		36	32	28	24	20	16	12	8	4
e_{o3}^{k}										
	2	0.0000	0.1531	0.1664	0.1971	0.3191	0.3665	0.4810	0.5882	0.6815
e_{o4}	1	0.0000	0.0855	0.0925	0.1100	0.1741	0.1998	0.2618	0.3260	0.3852
	2	0.0000	0.0855	0.0925	0.1100	0.1741	0.1998	0.2618	0.3260	0.3852
e_{o5}	1	0.0000	0.2165	0.2342	0.2787	0.4394	0.5045	0.6610	0.8252	0.9781
	2	0.0000	0.2165	0.2342	0.2787	0.4394	0.5045	0.6610	0.8252	0.9781

Table 4
Roughness measures of the decision class in the dataset Zoo.

Measure	Class	Number of condition attribute							
		16	14	12	10	8	6	4	2
ρ	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
	2	0.0000	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000
	3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.3333	0.3333	0.3846	0.9697	1.0000	1.0000
	5	0.0000	0.0000	0.4000	0.4000	0.4545	0.6000	0.8947	1.0000
	6	0.0000	0.0000	0.0000	0.0000	1.0000	1.0000	1.0000	1.0000
	7	0.0000	0.0000	0.0000	0.0000	0.5556	0.9000	1.0000	1.0000
E	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.7721
	2	0.0000	0.0000	0.0000	0.0000	0.4078	0.5535	0.7001	0.7721
	3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.1173	0.1275	0.1569	0.5367	0.7001	0.7721
	5	0.0000	0.0000	0.1407	0.1530	0.1854	0.3321	0.6264	0.7721
	6	0.0000	0.0000	0.0000	0.0000	0.4078	0.5535	0.7001	0.7721
	7	0.0000	0.0000	0.0000	0.0000	0.2266	0.4981	0.7001	0.7721
$e_{c1}^{D_M^1}$	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1188
	2	0.0000	0.0000	0.0000	0.0000	0.0792	0.1980	0.2574	0.2574
	3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.0396	0.0396	0.0396	0.1782	0.1980	0.1980
	5	0.0000	0.0000	0.0396	0.0396	0.0396	0.0396	0.0792	0.1584
	6	0.0000	0.0000	0.0000	0.0000	0.0792	0.0792	0.0792	0.0792
	7	0.0000	0.0000	0.0000	0.0000	0.0198	0.0792	0.0990	0.0990
$e_{nf1}^{D_M^2}$	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1375
	2	0.0000	0.0000	0.0000	0.0000	0.1940	0.1980	0.4437	0.4197
	3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.1990	0.1990	0.1780	0.1782	0.3379	0.3228
	5	0.0000	0.0000	0.1990	0.1990	0.1780	0.0396	0.1365	0.1772
	6	0.0000	0.0000	0.0000	0.0000	0.1940	0.0792	0.1365	0.1291
	7	0.0000	0.0000	0.0000	0.0000	0.0890	0.0792	0.1961	0.1614
$e_{o2}^{\alpha,\beta}$	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.3717
	2	0.0000	0.0000	0.0000	0.0000	0.1208	0.198	0.2722	0.2980
	3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.0328	0.0328	0.0403	0.1782	0.2669	0.2795
	5	0.0000	0.0000	0.0328	0.0328	0.0403	0.0396	0.1901	0.4226
	6	0.0000	0.0000	0.0000	0.0000	0.1208	0.0792	0.1901	0.2034
	7	0.0000	0.0000	0.0000	0.0000	0.0338	0.0792	0.2065	0.2217

(3) $\mu_{\widetilde{Y_B}}(x_p) \leqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \geqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \leqslant \frac{1}{2}$

$$\begin{aligned}\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) &= \mu_{\widetilde{Y_A}^F}(x_p) - \mu_{\widetilde{Y_B}^F}(x_p)\\ &= 1 - \mu_{\widetilde{Y_A}}(x_p) - \mu_{\widetilde{Y_B}}(x_p)\\ &= 1 - \frac{2|[x_p]_A \cap Y| \times |[x_p]_A| + |[x_p]_A \cap Y| \times |[x_q]_A| + |[x_p]_A| \times |[x_q]_A \cap Y|}{(|[x_p]_A| + |[x_q]_A|) \times |[x_p]_A|}.\end{aligned}$$

Similarly, we have

$$\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = \frac{|[x_q]_A \cap Y| \times |[x_p]_A| - |[x_q]_A| \times |[x_p]_A \cap Y|}{(|[x_q]_A| + |[x_p]_A|) \times |[x_q]_A|}.$$

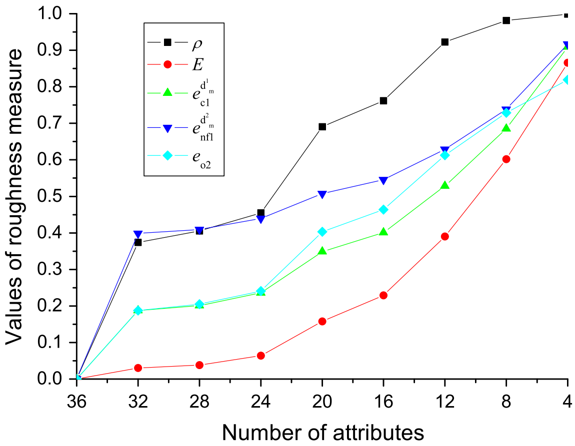
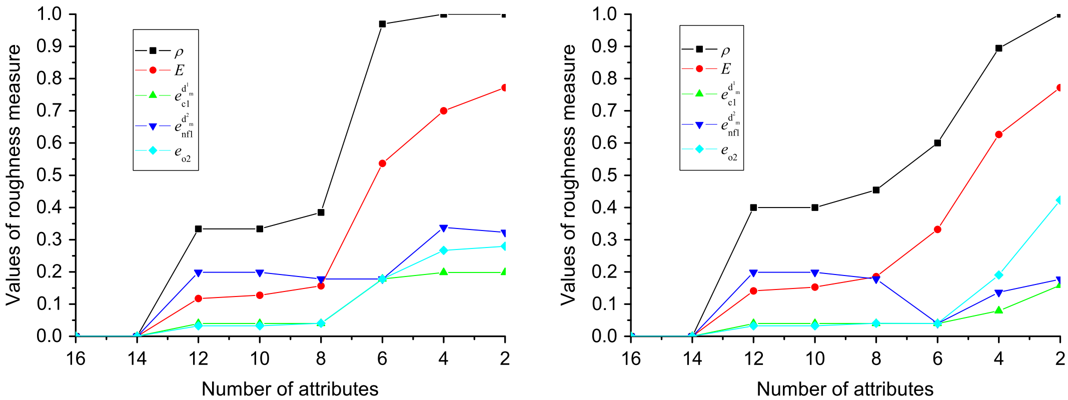


Fig. 3. Change of roughness measures in the dataset Kr-vs-kp.



(a) Rough set derived from Class 3

(b) Rough set derived from Class 4

Fig. 4. Change of roughness measures in the dataset Zoo.

Therefore,

$$|[x_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_q) = |[u_p]_A| - \frac{2|[x_p]_A \cap Y| \times |[x_p]_A| + 2|[x_p]_A \cap Y| \times |[x_q]_A|}{|[x_p]_A| + |[x_q]_A|}$$
$$= |[x_p]_A| - 2\mu_{\widetilde{Y_A}}(x_p) \times |[x_p]_A| \leqslant 0.$$

(4) $\mu_{\widetilde{Y_B}}(x_p) \leqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \leqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \geqslant \frac{1}{2}$

Because $[x_p]_A$ is symmetric to $[x_q]_A$ in $|[u_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_q)$, this case is analogous to Case (3). Thus,

$$|[x_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_q) \leqslant 0.$$

(5) $\mu_{\widetilde{Y_B}}(x_p) \geqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \leqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \geqslant \frac{1}{2}$

$$\Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_p) = \mu_{\widetilde{Y_A}^F}(x_p) - \mu_{\widetilde{Y_B}^F}(x_p) = \mu_{\widetilde{Y_A}}(x_p) - (1 - \mu_{\widetilde{Y_B}}(x_p)) = -\left(1 - \mu_{\widetilde{Y_A}}(x_p) - \mu_{\widetilde{Y_B}}(x_p)\right).$$

Similarly,

$$\Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_q) = -\left(\mu_{\widetilde{Y_A}}(x_q) - \mu_{\widetilde{Y_B}}(x_p)\right).$$

Similar to Case (3), we have

$$|[x_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F \widetilde{Y_A}^F}(x_q) = 2\mu_{\widetilde{Y_A}}(x_p) \times |[x_p]_A| - |[x_p]_A| \leqslant 0.$$

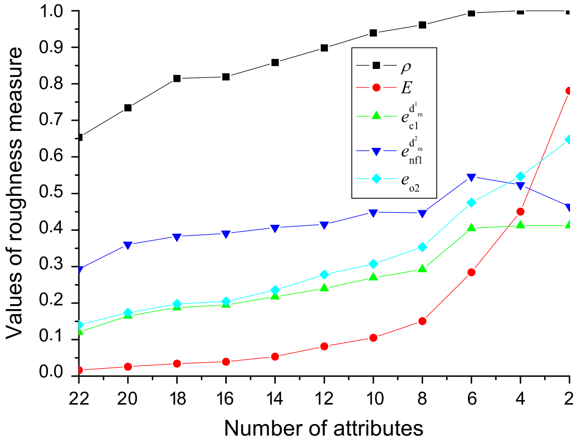


Fig. 2. Change of roughness measures in the dataset Spect.

(6) $\mu_{\widetilde{Y_B}}(x_p) \geqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \geqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \leqslant \frac{1}{2}$

Because the $[x_p]_A$ is symmetric to $[x_q]_A$ in $|[x_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)$, this case is similar to Case (5). Therefore, we have

$$|[x_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) \leqslant 0. \qquad \square$$

Theorem 3.1 states that if $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q) \neq \mu_{\widetilde{Y_B}}(x_p)$, then $|[x_p]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) \leqslant 0$. Because $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q)$, $\mu_{\widetilde{Y_B}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_p)$ and $\mu_{\widetilde{Y_B}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q)$ derive as $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q) \neq \mu_{\widetilde{Y_B}}(x_p)$ respectively, they are regarded as one case. Furthermore, it is obvious that $|[x_p]_B| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_B| \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = 0$ if $\mu_{\widetilde{Y_A}}(x_p) = \mu_{\widetilde{Y_A}}(x_q) = \mu_{\widetilde{Y_B}}(x_p)$.

In the following, we analyze the change of $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)$ with partition ordering.

Theorem 3.2. *Let $S_1 = (U, A)$ and $S_2 = (U, B)$ be two information systems, $Y \subseteq U$ and $\widetilde{Y_A}, \widetilde{Y_B} \in \mathbf{F}(U)$ two fuzzy sets. If $[x_p]_B = [x_p]_A \cup [x_q]_A$ $(x_p, x_q \in U)$, for $\forall x_i \notin [x_p]_B$ such that $[x_i]_B = [x_i]_A$, $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q) \neq \mu_{\widetilde{Y_B}}(x_p)$, and $\mu_{\widetilde{Y_B}}(x_p) \leqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \leqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \leqslant \frac{1}{2}$ (or $\mu_{\widetilde{Y_B}}(x_p) \geqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \geqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \geqslant \frac{1}{2}$), then*

$$\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) \leqslant 0,$$

where $\mu_{\widetilde{Y_A}}(x_i) = \frac{|[x_i]_A \cap Y|}{|[x_i]_A|}$, $\mu_{\widetilde{Y_B}}(x_i) = \frac{|[x_i]_B \cap Y|}{|[x_i]_B|}$, $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i) = \mu_{\widetilde{Y_A}^F}(x_i) - \mu_{\widetilde{Y_B}^F}(x_i)$.

Proof. (1) $\mu_{\widetilde{Y_B}}(x_p) \leqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \leqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \leqslant \frac{1}{2}$

$$\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = -\left(\frac{|[x_p]_A \cap Y| \times |[x_q]_A| - |[x_p]_A| \times |[x_q]_A \cap Y|}{(|[x_p]_A| + |[x_q]_A|) \times |[x_p]_A|}\right)^2 \leqslant 0.$$

By the existing condition $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q) \neq \mu_{\widetilde{Y_B}}(x_p)$, we have $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) < 0$.

(2) $\mu_{\widetilde{Y_B}}(x_p) \geqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \geqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \geqslant \frac{1}{2}$

$$\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = -\left(\frac{|[x_p]_A \cap Y| \times |[x_q]_A| - |[x_p]_A| \times |[x_q]_A \cap Y|}{(|[x_p]_A| + |[x_q]_A|) \times |[x_p]_A|}\right)^2 \leqslant 0.$$

By the existing condition $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q) \neq \mu_{\widetilde{Y_B}}(x_p)$, we have $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) < 0$. $\square$

Theorem 3.3. *Let $S_1 = (U, A)$ and $S_2 = (U, B)$ be two information systems, $Y \subseteq U$ and $\widetilde{Y_A}, \widetilde{Y_B} \in \mathbf{F}(U)$ two fuzzy sets. If $[x_p]_B = [x_p]_A \cup [x_q]_A$ $(x_p, x_q \in U)$ for $\forall x_i \notin [x_p]_B$ such that $[x_i]_B = [x_i]_A$, $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q) \neq \mu_{\widetilde{Y_B}}(x_p)$, and $\mu_{\widetilde{Y_B}}(x_p)$, $\mu_{\widetilde{Y_A}}(x_p)$ and $\mu_{\widetilde{Y_A}}(x_q)$ are not simultaneously greater or less than $\frac{1}{2}$, then the sign of $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)$ is uncertain, where $\mu_{\widetilde{Y_A}}(x_i) = \frac{|[x_i]_A \cap Y|}{|[x_i]_A|}$, $\mu_{\widetilde{Y_B}}(x_i) = \frac{|[x_i]_B \cap Y|}{|[x_i]_B|}$, $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i) = \mu_{\widetilde{Y_A}^F}(x_i) - \mu_{\widetilde{Y_B}^F}(x_i)$.*

Proof. To prove this theorem, four cases should be investigated as follows.

(1) $\mu_{\widetilde{Y_B}}(x_p) \leqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \geqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \leqslant \frac{1}{2}$

$$\mu_{\widetilde{Y_A}}(x_q) - \mu_{\widetilde{Y_B}}(x_p) = \frac{|[x_q]_A \cap Y| \times |[x_p]_A| - |[x_q]_A| \times |[x_p]_A \cap Y|}{(|[x_q]_A| + |[u_p]_A|) \times |[x_q]_A|} = \frac{\mu_{A_Y}(x_q) \times |[x_p]_A| - \mu_{A_Y}(x_p) \times |[x_p]_A|}{|[x_q]_A| + |[x_p]_A|} \leqslant 0.$$

Furthermore, it is obvious that the sign of $1 - \mu_{\widetilde{Y_A}}(x_p) - \mu_{\widetilde{Y_B}}(x_p)$ is uncertain by the condition $\mu_{\widetilde{Y_B}}(x_p) \leqslant \frac{1}{2}$ & $\mu_{\widetilde{Y_A}}(x_p) \geqslant \frac{1}{2}$ & $\mu_{\widetilde{Y_A}}(x_q) \leqslant \frac{1}{2}$.

Therefore, the sign of $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = (1 - \mu_{\widetilde{Y_A}}(x_p) - \mu_{\widetilde{Y_B}}(x_p)) \times (\mu_{\widetilde{Y_A}}(x_q) - \mu_{\widetilde{Y_B}}(x_p))$is also uncertain.

(2) $\mu_{\widetilde{Y_B}}(x_p) \leqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \leqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \geqslant \frac{1}{2}$

Because the $[x_p]_A$ is symmetric to $[x_q]_A$ in $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)$, this case is analogous to Case (1). Thus, we obtain that the sign of $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = (\mu_{\widetilde{Y_A}}(x_p) - \mu_{\widetilde{Y_B}}(x_p)) \times (1 - \mu_{\widetilde{Y_A}}(x_q) - \mu_{\widetilde{Y_B}}(x_p))$is also uncertain.

(3) $\mu_{\widetilde{Y_B}}(x_p) \geqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \leqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \geqslant \frac{1}{2}$

$$\begin{aligned}
&\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) \\
&= (\mu_{\widetilde{Y_A}}(x_p) - (1 - \mu_{\widetilde{Y_B}}(x_p))) \times (1 - \mu_{\widetilde{Y_B}}(x_p) - (1 - \mu_{\widetilde{Y_A}}(x_q))) \\
&= (1 - \mu_{\widetilde{Y_B}}(x_p) - \mu_{\widetilde{Y_A}}(x_p)) \times (\mu_{\widetilde{Y_A}}(x_q) - \mu_{\widetilde{Y_B}}(x_p)).
\end{aligned}$$

Similar to Case (1), the sign of $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)$ is also uncertain.

(4) $\mu_{\widetilde{Y_B}}(x_p) \geqslant \frac{1}{2}$, $\mu_{\widetilde{Y_A}}(x_p) \geqslant \frac{1}{2}$ and $\mu_{\widetilde{Y_A}}(x_q) \leqslant \frac{1}{2}$

Because $[x_p]_A$ is symmetric to $[x_q]_A$ in $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)$, this case is similar to Case (3). Therefore, the sign of $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = (1 - \mu_{\widetilde{Y_A}}(x_q) - \mu_{\widetilde{Y_B}}(x_p)) \times (\mu_{\widetilde{Y_A}}(x_p) - \mu_{\widetilde{Y_B}}(x_p))$is also uncertain. □

From Theorems 3.2 and 3.3, we can see that if $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q) \neq \mu_{\widetilde{Y_B}}(x_p)$, $\mu_{\widetilde{Y_A}}(x_p), \mu_{\widetilde{Y_A}}(x_q)$ and $\mu_{\widetilde{Y_B}}(x_p)$ are simultaneously greater or less than $\frac{1}{2}$ then $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) \leqslant 0$. Otherwise, the sign of $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)$ is uncertain. Because $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q), \mu_{\widetilde{Y_B}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_p)$ and $\mu_{\widetilde{Y_B}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q)$ deduce $\mu_{\widetilde{Y_A}}(x_p) \neq \mu_{\widetilde{Y_A}}(x_q) \neq \mu_{\widetilde{Y_B}}(x_p)$ respectively, they are regarded as one case. It is obvious that $\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q) = 0$ if $\mu_{\widetilde{Y_A}}(x_p) = \mu_{\widetilde{Y_A}}(x_q) = \mu_{\widetilde{Y_B}}(x_p)$.

To discover the general change mechanism of the fuzziness measure of a rough set, the fuzzy entropy is formulated using its Taylor's expansion. The following theorem indicates the corresponding results.

Theorem 3.4. *Let $S_1 = (U, A)$ and $S_2 = (U, B)$ be two information systems, $Y \subseteq U$ and $\widetilde{Y_A}, \widetilde{Y_B} \in \mathbf{F}(U)$ two fuzzy sets. If $[x_p]_B = [x_p]_A \cup [x_q]_A$ $(x_p, x_q \in U)$ and $[x_i]_B = [x_i]_A (\forall x_i \notin [x_p]_B)$, then*

$$\begin{aligned}
e(\widetilde{Y_A}) - e(\widetilde{Y_B}) &= f'_{\mu_{\widetilde{X}^F}(x_p)}(\mu_{\widetilde{Y_B}^F}) \times \left(|[x_p]_A|\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A|\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)\right) \\
&\quad + \frac{f''_{\mu_{\widetilde{X}^F}(u_p)\mu_{\widetilde{X}^F}(u_p)}(\mu_{\widetilde{Y_B}^F} + \xi)}{2} \times \left(|[x_p]_A|\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p)\right)^2 + |[x_q]_A|\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)\right)^2\right) \\
&\quad + f''_{\mu_{\widetilde{X}^F}(x_p)\mu_{\widetilde{X}^F}(x_q)}(\mu_{\widetilde{Y_B}^F} + \xi) \times |[x_p]_A| \times |[x_q]_A|\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)\right) \\
&\quad + \frac{f''_{\mu_{\widetilde{X}^F}(x_p)\mu_{\widetilde{X}^F}(x_q)}(\mu_{\widetilde{Y_B}^F} + \xi)}{2} \times (|[x_p]_A|^2 - |[x_p]_A|)\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p)\right)^2 \\
&\quad + \frac{f''_{\mu_{\widetilde{X}^F}(x_p)\mu_{\widetilde{X}^F}(x_q)}(\mu_{\widetilde{Y_B}^F} + \xi)}{2} \times (|[x_q]_A|^2 - |[x_q]_A|)\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)\right)^2,
\end{aligned}$$

Table 1
Description of three UCI datasets.

No.	Datasets	Objects	Condition attributes	Decision attributes	Classes
1	Spect	267	22	1	2
2	Kr-vs-kp	3196	36	1	2
3	Zoo	101	16	1	7

where $\Delta\mu_{\widetilde{Y_B}\widetilde{Y_A}}(x_i)=\mu_{\widetilde{Y_B}}(x_i)-\mu_{\widetilde{Y_A}}(x_i)$, $f'_{\mu_{\widetilde{X^F}}(x_i)}(\mu_{\widetilde{Y_B}^F})=f'_{\mu_{\widetilde{X^F}}(x_i)}(\mu_{\widetilde{Y_B}^F}(x_1),\mu_{\widetilde{Y_B}^F}(x_2),\ldots,\mu_{\widetilde{Y_B}^F}(x_n))$, $f''_{\mu_{\widetilde{X^F}}(x_i)\mu_{\widetilde{X^F}}(x_j)}(\mu_{\widetilde{Y_B}^F}+\xi)=f''_{\mu_{\widetilde{X^F}}(x_i)\mu_{\widetilde{X^F}}(x_j)}\left(\mu_{\widetilde{Y_B}^F}(x_1)+\xi_1\times\mu_{\widetilde{Y_B}^F}(x_1),\ldots,\mu_{\widetilde{Y_B}^F}(x_i)+\xi_i\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i),\ldots,\mu_{\widetilde{Y_B}^F}(x_n)+\xi_n\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_j)\right)$, *and* $\xi_i\in[0,1]$.

Proof. From Properties 3.1,3.2,3.3, we have $e(\widetilde{X})=e(\widetilde{X}^F)=f(\mu_{\widetilde{X^F}}(u_1),\mu_{\widetilde{X^F}}(u_2),\ldots,\mu_{\widetilde{X^F}}(u_n))$. Because $[x_p]_B=[x_p]_A\cup[x_q]_A$ for $x_p,x_q\in[x_p]_B$ and $[x_i]_B=[x_i]_A$ for $\forall x_i\notin[x_p]_B$, we have

$$\begin{aligned}
e(\widetilde{Y_A})-e(\widetilde{Y_B})&=e(\widetilde{Y_A}^F)-e(\widetilde{Y_B}^F)\\
&=\sum_{u_i\in U}f'_{\mu_{\widetilde{X^F}}(x_i)}(\mu_{\widetilde{Y_B}^F})\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)+\frac{1}{2}\sum_{x_i\in U}\sum_{x_j\in U}f''_{\mu_{\widetilde{X^F}}(x_i)\mu_{\widetilde{X^F}}(x_j)}(\mu_{\widetilde{Y_B}^F}+\xi_{ij})\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_j)\\
&=\sum_{x_i\in[x_p]_B}f'_{\mu_{\widetilde{X^F}}(x_i)}(\mu_{\widetilde{Y_B}^F})\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)+\frac{1}{2}\sum_{x_i\in[x_p]_B}\sum_{x_j\in[x_p]_B}\left(f''_{\mu_{\widetilde{X^F}}(x_i)\mu_{\widetilde{X^F}}(x_j)}(\mu_{\widetilde{Y_B}^F}+\xi)\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_j)\right).
\end{aligned}$$

Furthermore, because the function $e(\widetilde{X}^F)$ is symmetric with respect to every independent variable $\mu_{\widetilde{X^F}}(x_i)$, for $\forall x_i\in[x_p]_B$, one has that $f'_{\mu_{\widetilde{X^F}}(x_i)}(\mu_{\widetilde{X^F}})=f'_{\mu_{\widetilde{X^F}}(u_p)}(\mu_{\widetilde{Y_B}^F})$; for $\forall x_i,x_j\in[x_p]_B,\forall x_i\in[x_p]_A$ and $\forall x_j\in[x_q]_A$, one has that $f''_{\mu_{\widetilde{X^F}}(x_i)\mu_{\widetilde{X^F}}(x_i)}(\mu_{\widetilde{Y_B}^F}+\xi)=f''_{\mu_{\widetilde{X^F}}(x_p)\mu_{\widetilde{X^F}}(x_p)}(\mu_{\widetilde{Y_B}^F}+\xi)=f''_{\mu_{\widetilde{X^F}}(x_q)\mu_{\widetilde{X^F}}(x_q)}(\mu_{\widetilde{Y_B}^F}+\xi)=a$and one has that $f''_{\mu_{\widetilde{X^F}}(x_i)\mu_{\widetilde{X^F}}(x_j)}(\mu_{\widetilde{Y_B}^F}+\xi)=f''_{\mu_{\widetilde{X^F}}(x_j)\mu_{\widetilde{X^F}}(x_i)}(\mu_{\widetilde{Y_B}^F}+\xi)=f''_{\mu_{\widetilde{X^F}}(x_p)\mu_{\widetilde{X^F}}(x_q)}(\mu_{\widetilde{Y_B}^F}+\xi)=f''_{\mu_{\widetilde{X^F}}(x_q)\mu_{\widetilde{X^F}}(x_p)}(\mu_{\widetilde{Y_B}^F}+\xi)=b$, where $a,\ b$ are constants.

Therefore, we have

$$\begin{aligned}
e(\widetilde{Y_A})-e(\widetilde{Y_B})\ &=f'_{\mu_{\widetilde{X^F}}(x_p)}(\mu_{\widetilde{Y_B}^F})\times\left(\sum_{x_i\in[x_p]_A}\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)+\sum_{x_i\in[x_q]_A}\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)\right)\\
&\quad+\frac{a}{2}\times\left(\sum_{x_i\in[x_p]_A}\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)\right)^2+\sum_{x_i\in[x_q]_A}\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)\right)^2\right)\\
&\quad+\frac{b}{2}\times\sum_{x_i\in[x_p]_A}\sum_{x_j\in[x_q]_A}\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_j)\right)\\
&\quad+\frac{b}{2}\times\sum_{x_i\in[x_q]_A}\sum_{x_j\in[x_p]_A}\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_j)\right)\\
&\quad+\frac{b}{2}\times\sum_{x_i\in[x_p]_A}\sum_{x_j\in[x_p]_A,i\neq j}\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_j)\right)\\
&\quad+\frac{b}{2}\times\sum_{x_i\in[x_q]_A}\sum_{x_j\in[x_q]_A,i\neq j}\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i)\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_j)\right)\\
&=f'_{\mu_{\widetilde{X^F}}(x_p)}(\mu_{\widetilde{Y_B}^F})\times\left(|[x_p]_A|\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p)+|[x_q]_A|\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)\right)\\
&\quad+\frac{a}{2}\times\left(|[x_p]_A|\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p)\right)^2+|[x_q]_A|\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)\right)^2\right)\\
&\quad+b\times|[x_p]_A|\times|[x_q]_A|\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p)\times\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)\right)\\
&\quad+\frac{b}{2}\times(|[x_p]_A|^2-|[x_p]_A|)\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p)\right)^2\\
&\quad+\frac{b}{2}\times(|[x_q]_A|^2-|[x_q]_A|)\left(\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p)\right)^2.\qquad\square
\end{aligned}$$

Theorem 3.4 shows that the change of fuzzy entropies with the change of condition classes depends on the first- and second-order partial derivatives of a fuzzy entropy and the membership of each object. Thus, we introduce the following judging method.

Judging method 1 (JM1)

A fuzzy entropy can evaluate the roughness of a rough set if it satisfies the following conditions:

(1) its second-order partial derivative with respect to two different independent variables is equal to zero;

(2) its first-order partial derivative with respect to each independent variable is larger than zero;
(3) its second-order partial derivative with respect to the same independent variable is less than zero.

Remark. It is obvious that $\mu_{\widetilde{Y_A}^F}(x_i) = \mu_{\widetilde{Y_B}^F}(x_i)$ for $\forall x_i \in U$, i.e., $\Delta\mu_{\widetilde{Y_B}^F}\mu_{\widetilde{Y_A}^F} = \mu_{\widetilde{Y_B}^F}(x_i) - \mu_{\widetilde{Y_A}^F}(x_j) = 0$ if $BN_B(Y)/B = BN_A(Y)/A$ or if $|BN_B(Y)| = |BN_A(Y)|$, $BN_B(Y)/B \succ BN_A(Y)/A$ and $\mu_{\widetilde{Y_A}^F}(x_i) = \mu_{\widetilde{Y_B}^F}(x_i)$ for $\forall x_i \in BN_B(Y)$. Additionally, from Definition 3.1 and Theorem 3.4, we conclude that all of the fuzzy entropies satisfy (RP1) and (RP4). Furthermore, $\exists x_i \in BN_B(Y)$ such that $\mu_{\widetilde{Y_A}^F}(x_i) \neq \mu_{\widetilde{Y_B}^F}(x_i)$ if $|BN_B(Y)| > |BN_A(Y)|$ and $U/B \succ U/A$ or if $|BN_B(Y)| = |BN_A(Y)|, BN_B(Y)/B \succ BN_A(Y)/A$ and $\exists x_i \in BN_A(Y)$ such that $\mu_{\widetilde{Y_B}}(x_i) \neq \mu_{\widetilde{Y_A}}(x_i)$. Therefore, according to Definition 3.1 and Theorems 3.1,3.2,3.3,3.4, we have that the fuzzy entropies that fulfill the conditions in JM1 satisfy (RP2) and (RP3).

In the following, to provide the judging method for σ-entropies, we introduce the following theorem.

Theorem 3.5. *Let $e(\widetilde{X}) = f(\mu_{\widetilde{X}}(x_1), \mu_{\widetilde{X}}(x_2), \ldots, \mu_{\widetilde{X}}(x_n))$ and $\widetilde{X} \in \mathbf{F}(U)$. If e is a σ-entropy, then, for $\forall x_i \in U$,*

$$f'_{\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(x_n)) = f'_{\mu_{\widetilde{X}^F}(x_i)}(\underbrace{0, \ldots, 0}_{i-1}, \mu_{\widetilde{X}^F}(x_i), \underbrace{0, \ldots, 0}_{n-i}),$$

$$f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(x_n)) = f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_i)}(\underbrace{0, \ldots, 0}_{i-1}, \mu_{\widetilde{X}^F}(x_i), \underbrace{0, \ldots, 0}_{n-i}),$$

$$f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_j)}(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(x_n)) = 0.$$

Proof. Without any loss of generality, suppose $D = \{x_i\}$ and $D^c = \{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n\}$. By the conclusion of Theorem 2.2, $(e(\widetilde{X}) = e(\widetilde{X} \cap D) + e(\widetilde{X} \cap D^c))$, we have for $\forall x_i \in U$,

$$\begin{aligned} f'_{\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(u_n)) &= f'_{\mu_{\widetilde{X}^F}(x_i)}(\underbrace{0, \ldots, 0}_{i-1}, \mu_{\widetilde{X}^F}(x_i), \underbrace{0, \ldots, 0}_{n-i}) \\ &+ f'_{\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{X}^F}(x_1), \ldots, \mu_{\widetilde{X}^F}(x_{i-1}), 0, \mu_{\widetilde{X}^F}(x_{i+1}), \ldots, \mu_{\widetilde{X}^F}(x_n)) \\ &= f'_{\mu_{\widetilde{X}^F}(x_i)}(\underbrace{0, \ldots, 0}_{i-1}, \mu_{\widetilde{X}^F}(x_i), \underbrace{0, \ldots, 0}_{n-i}). \quad \square \end{aligned}$$

Theorem 3.5 states that for σ-entropy, the first condition in JM1 is naturally satisfied. Therefore, a σ-entropy can be used to evaluate the roughness of a rough set if the σ-entropy satisfies the last two conditions in JM1.

Furthermore, by the results of Theorems 3.2 and 3.3, the change of fuzzy entropies with the combination of condition classes results in uncertainty when the second-order partial derivative with respect to two different independent variables is not equal to zero. According to the above analysis, we present another method as follows.

Judging method 2 (JM2)

A fuzzy entropy cannot evaluate the roughness of a rough set if the fuzzy entropy satisfies the following condition: its second-order partial derivative with respect to two different independent variable is not zero.

Remark. From Definition 3.1 and Theorems 3.2,3.3,3.4, we have that the fuzzy entropies that satisfy the condition in JM2 are contrary to (RP2) and (RP4).

To investigate other methods, we provide the following corollary, which can be straightforwardly derived from Theorem 3.4.

Corollary 3.1. *Let $S_1 = (U, A)$ and $S_2 = (U, B)$ be two information systems, $Y \subseteq U$ and $\widetilde{Y_A}, \widetilde{Y_B} \in \mathbf{F}(U)$ two fuzzy sets. If $[x_p]_B = [x_p]_A \cup [x_q]_A$ $(x_p, x_q \in U)$, for $\forall x_i \notin [x_p]_B$ such that $[x_i]_B = [x_i]_A$ and $f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{Y_B}^F} + \xi) = f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_j)}(\mu_{\widetilde{Y_B}^F} + \xi), (i \neq j)$, then*

$$\begin{aligned} e(\widetilde{Y_A}) - e(\widetilde{Y_B}) = &f'_{\mu_{\widetilde{X}^F}(x_p)}(\mu_{\widetilde{Y_B}^F}) \times \left(|[x_p]_A|\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A|\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)\right) + \frac{f''_{\mu_{\widetilde{X}^F}(x_p)\mu_{\widetilde{X}^F}(x_p)}(\mu_{\widetilde{Y_B}^F} + \xi)}{2} \\ &\times \left(|[x_p]_A|\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_p) + |[x_q]_A|\Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_q)\right)^2, \end{aligned}$$

where $\Delta\mu_{\widetilde{Y_B}\widetilde{Y_A}}(x_i) = \mu_{\widetilde{Y_B}}(x_i) - \mu_{\widetilde{Y_A}}(x_i)$, $f'_{\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{Y_B}^F}) = f'_{\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{Y_B}^F}(x_1), \mu_{\widetilde{Y_B}^F}(x_2), \ldots, \mu_{\widetilde{Y_B}^F}(x_n))$, $f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_j)}(\mu_{\widetilde{Y_B}^F} + \xi) = f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_j)}\Big(\mu_{\widetilde{Y_B}^F}(x_1) + \xi_1 \times \mu_{\widetilde{Y_B}^F}(x_1), \ldots, \mu_{\widetilde{Y_B}^F}(x_i) + \xi_i \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_i), \ldots, \mu_{\widetilde{Y_B}^F}(x_n) + \xi_n \times \Delta\mu_{\widetilde{Y_B}^F\widetilde{Y_A}^F}(x_j)\Big)$, *and* $\xi_i \in [0, 1]$.

From Corollary 3.1, we can see that for the fuzzy sets derived from the two rough sets that are approximated by two ordering partitions, the fuzzy entropy values are equal when the membership degrees of each object in the fuzzy sets are simultaneously greater and less than $\frac{1}{2}$ and their second-order partial derivatives with respect to two different independent variables are equal to the second-order partial derivative with respect to the same independent variable. From this corollary, we propose Judging method 3.

Judging method 3 (JM3)

A fuzzy entropy cannot evaluate the roughness of a rough set if the fuzzy entropy satisfies the following condition: its second-order partial derivative with respect to two different independent variables is equal to the second-order partial derivative with respect to the same independent variable.

Remark. From the Definition 3.1 and Theorems 3.1 and 3.4, we obtain that the fuzzy entropies that satisfy the condition in JM3 do not meet (RP2) and (RP4).

4. Using fuzzy entropies to evaluate the roughness of rough sets

Because a rough set can induce a fuzzy set when an object's rough membership degree is regarded as its fuzzy membership degree. Therefore, we speculate that certain fuzzy entropies can characterize the roughness of a rough set. In this section, we investigate which fuzzy entropies can be used to evaluate roughness using the judging methods introduced above.

4.1. Evaluating roughness with a σ-entropy

For convenience, σ-entropies are divided into three types: fuzzy entropies based on the Minkowski distance ($p = 1$), fuzzy entropies based on D_E and D_L and other entropies.

The σ-entropies based on the Minkowski distance ($p = 1$)

Klir et al. [17] proposed the fuzzy Minkowski distance measure between two fuzzy sets. The fuzzy Minkowski distance is σ-distance when the parameter $p = 1$. In the following, we analyze the σ-entropies based on this distance.

Proposition 4.1. *The fuzzy entropies $e_{c1}^{D_M^1}$, $e_{c6}^{D_M^1}$, $e_{nf1}^{D_M^1}$, $e_{nf4}^{D_M^1}$, $e_{o1}^{D_M^1}$ cannot evaluate the roughness of a rough set.*

Proof. By Property 3.1 and the definition of the near set, we have

$$e_{nf1}^{D_M^1}(\widetilde{X}^F) = f(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(x_n)) = \frac{2}{n}\sum_{i=1}^{n}\mu_{\widetilde{X}^F}(x_i).$$

By calculation,

$$f'_{\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(x_n)) = \frac{2}{n} > 0, \quad \forall x_i \in U,$$

$$f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(x_n)) = 0, \quad \forall x_i \in U,$$

$$f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_j)}(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(x_n)) = 0, \quad \forall x_i, x_j \in U,\ i \neq j.$$

It is obvious that $f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_i)}(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(x_n)) = f''_{\mu_{\widetilde{X}^F}(x_i)\mu_{\widetilde{X}^F}(x_j)}(\mu_{\widetilde{X}^F}(x_1), \mu_{\widetilde{X}^F}(x_2), \ldots, \mu_{\widetilde{X}^F}(x_n))(i \neq j)$. Therefore, using the method JM3, we obtain that $e_{c1}^{D_M^1}, e_{c6}^{D_M^1}, e_{nf1}^{D_M^1}, e_{nf4}^{D_M^1}, e_{o1}^{D_M^1}$ cannot evaluate the roughness of a rough set. □

The following example illustrates the results of the above proposition.

Example 4.1. Let $S_1 = (U, A), S_2 = (U, B)$, $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$, $U/A = \{\{x_1, x_2, x_3, x_4, x_5\}, \{x_6, x_7, x_8\}, \{x_9, x_{10}\}\}$, $U/B = \{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}, \{x_9, x_{10}\}\}$, and $Y = \{x_1, x_4, x_7, x_9, x_{10}\}$. By calculation, we obtain $BN_A(Y) = \{x_1, x_2, x_3, x_4, x_5\} \cup \{x_6, x_7, x_8\}$ and $BN_B(Y) = \{x_1, x_2, x_3, u_4, x_5, x_6, x_7, x_8\}$.

It is evident that $BN_B(Y)/B \succ BN_A(Y)/A$ *and* $|BN_A(Y)| = |BN_B(Y)|$. Furthermore, the two fuzzy sets $\widetilde{Y_A}, \widetilde{Y_B} \in \mathbf{F}(U)$ can be constructed by the two rough sets $(\overline{B}(Y), \underline{B}(Y))$ and $(\overline{A}(Y), \underline{A}(Y))$, respectively, which are denoted as

$$\widetilde{Y_A} = \frac{2/5}{x_1} + \frac{2/5}{x_2} + \frac{2/5}{x_3} + \frac{2/5}{x_4} + \frac{2/5}{x_5} + \frac{1/3}{x_6} + \frac{1/3}{x_7} + \frac{1/3}{x_8} + \frac{1}{x_9} + \frac{1}{x_{10}},$$

$$\widetilde{Y_B} = \frac{3/8}{x_1} + \frac{3/8}{x_2} + \frac{3/8}{x_3} + \frac{3/8}{x_4} + \frac{3/8}{x_5} + \frac{3/8}{x_6} + \frac{3/8}{x_7} + \frac{3/8}{x_8} + \frac{1}{x_9} + \frac{1}{x_{10}}.$$

Therefore, $\mu_{\widetilde{Y_A}}(x_1) \neq \mu_{\widetilde{Y_B}}(x_1),\ \mu_{\widetilde{Y_A}}(x_2) \neq \mu_{\widetilde{Y_B}}(x_2) \neq \cdots \neq \mu_{\widetilde{Y_A}}(x_8) \neq \mu_{\widetilde{Y_B}}(x_8),\ \mu_{\widetilde{Y_A}}(x_9) = \mu_{\widetilde{Y_B}}(x_9),\ \mu_{\widetilde{Y_A}}(x_{10}) = \mu_{\widetilde{Y_B}}(x_{10})$.

By calculation, one obtains $e_{c1}^{D_M^1}(\widetilde{Y_A}) = e_{c1}^{D_M^1}(\widetilde{Y_B}) = 0.6000$, $e_{c6}^{D_M^1}(\widetilde{Y_A}) = e_{c6}^{D_M^1}(\widetilde{Y_B}) = 0.6000$, $e_{nf1}^{D_M^1}(\widetilde{Y_A}) = e_{nf1}^{D_M^1}(\widetilde{Y_B}) = 0.6000$, $e_{nf4}^{D_M^1}(\widetilde{Y_A}) = e_{nf4}^{D_M^1}(\widetilde{Y_B}) = 0.6000$, $e_{o1}^{D_M^1}(\widetilde{Y_A}) = e_{o1}^{D_M^1}(\widetilde{Y_B}) = 0.6000$, which contradicts (RP3).

The σ-entropies based on D_E and D_L

Bhandari et al. [3] introduced a distance D_L, and Fan and Xie [8] proposed another distance D_E. The two distance measures are both σ-distances. The σ-entropies based on these distances are analyzed as follows.

Proposition 4.2. *The fuzzy entropies $e_{c1}^{D_E},\ e_{c1}^{D_L},\ e_{c6}^{D_E},\ e_{c6}^{D_L},\ e_{nf4}^{D_E}$ and $e_{nf4}^{D_L}$ is able to evaluate the roughness of a rough set.*

By the method JM1, we straightforwardly prove the above proposition. The following example illustrates its results.

Example 4.2 (*Continued from Example 2.3*). $\widetilde{Y_B}, \widetilde{Y_{A_1}}, \widetilde{Y_{A_2}}, \widetilde{Y_{A_3}}, \widetilde{Y_{A_4}} \in \mathbf{F}(U)$ are five fuzzy sets derived from rough sets $(\overline{B}(Y), \underline{B}(Y))$, $(\overline{A_1}(Y), \underline{A_1}(Y))$, $(\overline{A_2}(Y), \underline{A_2}(Y))$, $(\overline{A_3}(Y), \underline{A_3}(Y))$ and $(\overline{A_4}(Y), \underline{A_4}(Y))$.

By calculation, one obtains

$$e_{c1}^{D_E}(\widetilde{Y_B}) = 0.9548,\quad e_{c1}^{D_E}(\widetilde{Y_{A_1}}) = 0.7181,\quad e_{c1}^{D_E}(\widetilde{Y_{A_2}}) = 0.6931,\quad e_{c1}^{D_E}(\widetilde{Y_{A_3}}) = 0.7181,\quad e_{c1}^{D_E}(\widetilde{Y_{A_4}}) = 0.7181,$$

$$e_{c1}^{D_L}(\widetilde{Y_B}) = 0.9463,\quad e_{c1}^{D_L}(\widetilde{Y_{A_1}}) = 0.6988,\quad e_{c1}^{D_L}(\widetilde{Y_{A_2}}) = 0.6715,\quad e_{c1}^{D_L}(\widetilde{Y_{A_3}}) = 0.6988,\quad e_{c1}^{D_L}(\widetilde{Y_{A_4}}) = 0.6988,$$

$$e_{c6}^{D_E}(\widetilde{Y_B}) = 0.8494,\quad e_{c6}^{D_E}(\widetilde{Y_{A_1}}) = 0.5296,\quad e_{c6}^{D_E}(\widetilde{Y_{A_2}}) = 0.5272,\quad e_{c6}^{D_E}(\widetilde{Y_{A_3}}) = 0.5296,\quad e_{c6}^{D_E}(\widetilde{Y_{A_4}}) = 0.5296,$$

$$e_{c6}^{D_L}(\widetilde{Y_B}) = 0.8362,\quad e_{c6}^{D_L}(\widetilde{Y_{A_1}}) = 0.5053,\quad e_{c6}^{D_L}(\widetilde{Y_{A_2}}) = 0.5049,\quad e_{c6}^{D_L}(\widetilde{Y_{A_3}}) = 0.5053,\quad e_{c6}^{D_L}(\widetilde{Y_{A_4}}) = 0.5053,$$

$$e_{nf4}^{D_E}(\widetilde{Y_B}) = 0.8494,\quad e_{nf4}^{D_E}(\widetilde{Y_{A_1}}) = 0.5296,\quad e_{nf4}^{D_E}(\widetilde{Y_{A_2}}) = 0.5272,\quad e_{nf4}^{D_E}(\widetilde{Y_{A_3}}) = 0.5296,\quad e_{nf4}^{D_E}(\widetilde{Y_{A_4}}) = 0.5296,$$

$$e_{nf4}^{D_L}(\widetilde{Y_B}) = 0.8494,\quad e_{nf4}^{D_L}(\widetilde{Y_{A_1}}) = 0.5296,\quad e_{nf4}^{D_L}(\widetilde{Y_{A_2}}) = 0.5272,\quad e_{nf4}^{D_L}(\widetilde{Y_{A_3}}) = 0.5296,\quad e_{nf4}^{D_L}(\widetilde{Y_{A_4}}) = 0.5296.$$

Obviously, we have

(1) $e_{c1}^{D_E}(\widetilde{Y_{A_1}}) = e_{c1}^{D_E}(\widetilde{Y_{A_3}})$ when $BN_{A_1}(Y)/A_1 = BN_{A_3}(Y)/A_3$ & $|BN_{A_1}(Y)/A_1| = |BN_{A_3}(Y)/A_3|$, which is according with (RP1);

(2) $e_{c1}^{D_E}(\widetilde{Y_B}) > e_{c1}^{D_E}(\widetilde{Y_{A_1}}),\ e_{c1}^{D_E}(\widetilde{Y_B}) > e_{c1}^{D_E}(\widetilde{Y_{A_2}}),\ e_{c1}^{D_E}\widetilde{Y_B} > e_{c1}^{D_E}(Y_{A_3}),\ e_{c1}^{D_E}(\widetilde{Y_B}) > e_{c1}^{D_E}(\widetilde{Y_{A_4}})$ when $U/B \succ U/A_1$ & $|BN_B(Y)/B| > |BN_{A_1}(Y)/A_1|,\ U/B \succ U/A_2$ & $|BN_B(Y)/B| > |BN_{A_2}(Y)/A_2|,\ U/B \succ U/A_3$ & $|BN_B(Y)/B| > |BN_{A_3}(Y)/A_3|$ and $U/B \succ U/A_4$ & $|BN_B(Y)/B| > |BN_{A_4}(Y)/A_4|$, respectively, which are coincident with (RP2);

(3) $e_{c1}^{D_E}(\widetilde{Y_{A_1}}) > e_{c1}^{D_E}(\widetilde{Y_{A_2}})$ when $BN_{A_1}(Y)/A_1 \succ BN_{A_2}(Y)/A_2$ & $|BN_{A_1}(Y)/A_1| = |BN_{A_2}(Y)/A_2|$ & $\mu_{A_1}(x_1) = \mu_{A_1}(x_2) < \mu_{A_2}(x_1) = \mu_{A_2}(x_2)$, which satisfies (RP3);

(4) $e_{c1}^{D_E}(\widetilde{Y_{A_1}}) = e_{c1}^{D_E}(\widetilde{Y_{A_4}})$ when $BN_{A_1}(Y)/A_1 \succ BN_{A_2}(Y)/A_2$ & $|BN_{A_1}(Y)/A_1| = |BN_{A_2}(Y)/A_2|$ & $\mu_{\widetilde{Y_{A_1}}}(x_1) = \cdots = \mu_{\widetilde{Y_{A_1}}}(x_{12}) = \mu_{\widetilde{Y_{A_4}}}(x_1) = \cdots = \mu_{\widetilde{Y_{A_4}}}(x_{12})$, which corresponds with (RP4).

Similarly, it is obvious that $e_{c1}^{D_E}, e_{c1}^{D_L},\ e_{c6}^{D_E},\ e_{c6}^{D_L}, e_{nf4}^{D_E}$ and $e_{nf4}^{D_L}$ satisfy (RP1)–(RP4).

Other σ-entropies

In addition to the σ-entropies induced from distance, there are several common entropies. In this section, we will investigate whether the common entropies can be used to evaluate the roughness of a rough set.

Proposition 4.3. *The fuzzy entropies $e_{o2}^{\alpha,\beta},\ e_{o3}^{k},\ e_{o4},\ e_{o5}$ is capable to evaluate the roughness of a roughness.*

It is straightforward to prove the above proposition by the method JM1, which is illustrated by Example 4.3.

Example 4.3. (Continued from Example 2.3) Without any loss of generality, let the parameter $k = 1,\ \alpha = 0.5$ and $\beta = 1$. By calculation, we obtain

$$e_{o2}^{\alpha,\beta}(\widetilde{Y_B}) = 0.8081,\quad e_{o2}^{\alpha,\beta}(\widetilde{Y_{A_1}}) = 0.6377,\quad e_{o2}^{\alpha,\beta}(\widetilde{Y_{A_2}}) = 0.6261,\quad e_{o2}^{\alpha,\beta}(\widetilde{Y_{A_3}}) = 0.6377,\quad e_{o2}^{\alpha,\beta}(\widetilde{Y_{A_4}}) = 0.6377,$$

$$e_{o3}^{k}(\widetilde{Y_B}) = 0.6648,\quad e_{o3}^{k}(\widetilde{Y_{A_1}}) = 0.5057,\quad e_{o3}^{k}(\widetilde{Y_{A_2}}) = 0.4904,\quad e_{o3}^{k}(\widetilde{Y_{A_3}}) = 0.5057,\quad e_{o3}^{k}(\widetilde{Y_{A_4}}) = 0.5057,$$

$$e_{o4}(\widetilde{Y_B}) = 0.7874,\quad e_{o4}(\widetilde{Y_{A_1}}) = 0.5799,\quad e_{o4}(\widetilde{Y_{A_2}}) = 0.5574,\quad e_{o4}(\widetilde{Y_{A_3}}) = 0.5799,\quad e_{o4}(\widetilde{Y_{A_4}}) = 0.5799,$$

$$e_{o5}(\widetilde{Y_B}) = 0.9444,\quad e_{o5}(\widetilde{Y_{A_1}}) = 0.6944,\quad e_{o5}(\widetilde{Y_{A_2}}) = 0.6667,\quad e_{o5}(\widetilde{Y_{A_3}}) = 0.6944,\quad e_{o5}(\widetilde{Y_{A_4}}) = 0.6944.$$

As in Example 2.3, it is obvious that $e_{o2}^{\alpha,\beta}, e_{o3}^{k}, e_{o4}ande_{o5}$ can evaluate the roughness of a rough set.

4.2. Evaluating roughness with non-σ-entropy

In this section, for convenience, non-σ-entropies are divided into three types: fuzzy entropies based on the Minkowski distance ($p = 1$), fuzzy entropies based on the Minkowski distance ($p \geqslant 2$) and fuzzy entropies based on D_E and D_L.

$e_{c2}^{D_M^1}$ and e_{nf2} are constructed based on the Minkowski distance ($p = 1$). We first analyze whether they can be used to evaluate the roughness of a rough set.

Proposition 4.4. *The fuzzy entropies* $e_{c2}^{D_M^1}$, e_{nf2} *is incapable to evaluate the roughness of a rough set.*

It is straightforward to prove Proposition 4.4 by JM3. The following example illustrates the proposition.

Example 4.4 (*Continued from Example 4.1*). By calculation, one obtains $e_{c2}^{D_M^1}(\widetilde{Y_A}) = e_{c2}^{D_M^1}(\widetilde{Y_B}) = 0.4286$, $e_{nf2}(\widetilde{Y_A}) = e_{nf2}(\widetilde{Y_B}) = 0.4286$, which contracts with (RP3).

Proposition 4.5. *The fuzzy entropies* $e_{c2}^{D_M^p}$, $e_{c6}^{D_M^p}$ *and* $e_{nf1}^{D_M^p}$ *cannot evaluate the roughness of a rough set.*

This proposition is straightforward to be proved with JM2, which is illustrated by the following example.

Example 4.5. (*Continued from Example 4.1*). For simplicity, without any loss of generality, we designate the parameter p as 2. By calculation, we obtain $e_{c2}^2(\widetilde{Y_A}) = 0.4699$, $e_{nf1}^2(\widetilde{Y_B}) = 0.4685$, $e_{nf1}^2(\widetilde{Y_A}) = 0.6202$, $e_{nf1}^2(\widetilde{Y_B}) = 0.6195$, $e_{nf1}^2(\widetilde{Y_A}) = 0.6733ande_{nf1}^2(\widetilde{Y_B}) = 0.6708$. It is obvious that $e_{c2}^{D_M^2}(\widetilde{Y_B}) < e_{c2}^{D_M^2}(\widetilde{Y_A}), e_{c6}^{D_M^2}(\widetilde{Y_B}) < e_{c6}^{D_M^2}(\widetilde{Y_A}), e_{nf1}^{D_M^2}(\widetilde{Y_B}) < e_{nf1}^{D_M^2}(\widetilde{Y_A})$, which does not accord with (RP3).

Bhandari et al. [3] introduced a distance D_L, and Fan and Xie [8] proposed another distance D_E. The two distance measures are both σ-distances. Based on these measures, many fuzzy entropies were proposed. In the following, We investigate whether they can be used to evaluate the roughness of a rough set.

Proposition 4.6. *The fuzzy entropies* $e_{c2}^{D_E}$, $e_{c2}^{D_L}$, $e_{nf3}^{D_E}$ *and* $e_{nf3}^{D_L}$ *cannot evaluate the roughness of a rough set.*

It is straightforward to prove the proposition by JM2. The following example illustrates its results.

Example 4.6. (*Continued from Example 4.1*). By calculation, we obtain $e_{c2}^{D_E}(\widetilde{Y_A}) = 0.1965$, $e_{c2}^{D_E}(\widetilde{Y_B}) = 0.1953$, $e_{nf3}^{D_L}(\widetilde{Y_A}) = 0.2243$, $e_{nf3}^{D_L}(\widetilde{Y_B}) = 0.2231$, $e_{nf3}^{D_E}(\widetilde{Y_A}) = 0.1965$, $e_{nf3}^{D_E}(\widetilde{Y_B}) = 0.1953$, $e_{nf3}^{D_L}(\widetilde{Y_A}) = 0.2243$ and $e_{nf3}^{D_L}(\widetilde{Y_B}) = 0.2231$. It is obvious that $e_{c2}^{D_E}(\widetilde{Y_B}) < e_{c2}^{D_E}(\widetilde{Y_A})$, $e_{c2}^{D_L}(\widetilde{Y_B}) < e_{c2}^{D_L}(\widetilde{Y_A})$, $e_{nf3}^{D_E}(\widetilde{Y_B}) < e_{nf3}^{D_E}(\widetilde{Y_A})$ and $e_{nf3}^{D_L}(\widetilde{Y_B}) < e_{nf3}^{D_L}(\widetilde{Y_A})$, which is incompatible with (RP3).

Finally, the method proposed by us is not appropriate for certain entropies, such as $e_{c1}^{D_M^p}(p \geqslant 2)$, because their second-order partial derivatives with two different independent variables (the membership degree of each object) are not zero.

5. Experiment analysis

In this section, we will analyze how Pawlak's roughness measure, Beaubouef's rough entropy and the fuzzy entropies change with the number of attributes provided in the practical datasets, which illustrate the results of Section 4 from the perspective of experiments. The three datasets used in the experiment, described in Table 1, come from the UCI Machine Learning Repository. All of the measures in this paper were computed on a personal computer equipped with Intel Core2 Quad CPU Q9400 and 2 GB Memory. The operation system is Windows XP. In the following tables and figures, all of these measures have been normalized.

5.1. The analysis of Pawlak's roughness measure and Beaubouef's rough entropy

In the datasets Spect and Kr-vs-kp, there are two decision classes. They are obviously complementary, and the rough set generated by using the condition attributes to approximate them are identical in size and structure. Therefore, Pawlak's roughness measure and Beaubouef's rough entropy values for the two decision classes should be equal. However, from Table 2, Table 3, we observe that their values for the two decision classes are different when the condition attributes are same. In the dataset Zoo, there are seven decision classes. Therefore, the rough sets derived from the seven decision classes are different. Thus, the Pawlak's roughness measure and Beaubouef's rough entropy values for the seven decision classes are not necessarily equal. Table 4 shows the results.

In Table 2, Pawlak's roughness measure for Class 2 is invariant when the number of attributes decreases to 2 from 4. A similar result is presented in Table 4. The similarity occurs because the boundary region and upper approximation of each class are unchanged. Furthermore, as shown in Figs. 2–4, the value range of Beaubouef's rough entropy is significantly wider than that of the fuzzy entropies and is comparable with Pawlak's roughness measure, which is caused by superfluously including the granule change of the lower and upper approximations.

5.2. The analysis of fuzzy entropies

From Table 2, we see the changes of the fuzzy entropies with the attributes from the dataset Spect. Thus, the fuzzy entropies are divided into three types: (1) strictly increasing ($e_{c1}^{D_M^2}$, $e_{c1}^{D_E}$, $e_{c1}^{D_L}$, $e_{c6}^{D_E}$, $e_{c6}^{D_L}$, $e_{nf4}^{D_E}$, $e_{nf4}^{D_L}$, $e_{o2}^{\alpha,\beta}$, e_{o3}^{k}, e_{o4} and e_{o5}); (2) non-decreasing ($e_{c1}^{D_M^1}$, $e_{c6}^{D_M^1}$, $e_{c2}^{D_M^1}$, $e_{nf1}^{D_M^1}$, e_{nf2}, $e_{nf4}^{D_M^1}$ and $e_{o1}^{D_M^1}$); and (3) non-monotone ($e_{c2}^{D_M^2}$, $e_{c6}^{D_M^2}$, $e_{nf1}^{D_M^2}$, $e_{c2}^{D_E}$, $e_{c2}^{D_L}$, $e_{nf3}^{D_E}$ and $e_{nf3}^{D_L}$).

Furthermore, from Table 3, we obtain that all types of the fuzzy entropies are strictly increasing with the number of attributes in the dataset Kr-vs-kp. Based on the results provided in Tables 2 and 3, it is clear that the first type of fuzzy entropy is strictly increasing in the two datasets that are suitable for evaluating the roughness of a rough set. This conclusion comply with Section 4. To illustrate this conclusion, we present in Figs. 2,3 a representative from each type of fuzzy entropy.

The experimental results on the dataset Zoo with seven decision classes are shown in Table 4 and Fig. 4. For brevity, only three representative entropies are analyzed. Figs. 4a and b illustrate the changes of the roughness measures of the rough sets derived from Classes 3 and 4, respectively. Fig. 4 shows that these fuzzy entropies change in accordance with the datasets Spect and Kr-vs-kp.

6. Conclusion

In this paper, we have discussed limitations of the existing roughness measures and introduced a more effective roughness measure of a rough set. Three methods have been proposed for determining whether fuzzy entropies can be measures for evaluating the roughness of a rough set. The applications of these methods have led to the following conclusions:

(1) For σ-entropy, if its first partial derivative with respect to every independent variable is larger than zero and its second-order partial derivative with respect to the same independent variable is less than zero, it can evaluate the roughness of a rough set.

(2) For non-σ-entropy, if its second partial derivative with respect to two different independent variables is equal to the second partial derivative with respect to the same independent variables, or if its second-order partial derivative with respect to two different independent variables is not zero, it cannot evaluate the roughness of a rough set.

It is our hope that this paper could provide some new ideas on evaluating the roughness of a rough set. Our further research will focus on applications of the proposed methods in testing whether fuzziness measures can evaluate the roughness in an incomplete rough set model [4], a fuzzy rough set model [11,12], a probability rough set model [55] and a neighborhood rough set model [14,56]. The results of this paper for the integration of roughness and fuzziness measures can also be utilized in image-processing [27,28,34], clustering [52] and feature selection [42,46].

Acknowledgements

The authors are grateful to the editors and the anonymous referees for their thoughtful comments and valuable suggestions, which helped us improve this paper. The research was supported by the National Natural Science Foundation of China (Nos. 71031006, 70971080, 60903110, and 60970014), the Special Prophase Project for the National Key Basic Research and Development Program of China (973) (No. 2011CB311805), and the Natural Science Foundation of Shanxi Province (Nos. 2010021017-3 and 2009021017-1).

References

[1] M. Banerjee, S.K. Pal, Roughness of a fuzzy set, Information Sciences 93 (1996) 235–246.
[2] T. Beaubouef, F.E. Petry, G. Arora, Information-theoretic measures of uncertainty for rough sets and rough relational databases, Information Sciences 109 (1998) 535–563.
[3] D. Bhandari, N.R. Pal, Some new information measure of fuzzy sets, Information Sciences 67 (1993) 209–228.
[4] D. Bianucci, G. Cattaneo, Information entropy and granulation co-entropy of partitions and coverings: a summary, in: J.F. Peters et al. (Eds.), Transactions on Rough Sets 10, LNCS 5656, 2009, pp. 15–66.
[5] K. Chakrabarty, R. Biswas, S. Nanda, Fuzziness in rough sets, Fuzzy Sets and Systems 110 (2000) 247–251.
[6] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems 17 (1990) 191–208.
[7] D. Dubois, H. Prade, Putting fuzzy sets and rough sets together, in: R. Slowiniski (Ed.), Intelligent Decision Support, Kluwer Academic, Dordrecht, 1992, pp. 203–232.
[8] J.L. Fan, W.X. Xie, Distance measure and induced fuzzy entropy, Fuzzy Sets and Systems 104 (1999) 305–314.
[9] J.L. Fan, Y.L. Ma, W.X. Xie, On some properties of distance measures, Fuzzy Sets and Systems 117 (2001) 355–361.
[10] J.L. Fan, Y.L. Ma, Some new fuzzy entropy formulas, Fuzzy Sets and Systems 128 (2002) 277–284.
[11] Q.H. Hu, Z.X. Xie, D.R. Yu, Fuzzy probabilistic approximation spaces and their information measures, IEEE Transactions on Fuzzy Systems 14 (2006) 191–201.
[12] Q.H. Hu, D.R. Yu, Z.X. Xie, Information-preserving hybrid data reduction based on fuzzy-rough techniques, Pattern Recognition Letters 27 (2006) 414–423.
[13] Q.H. Hu, Z.X. Xie, D.R. Yu, Hybrid attribute reduction based on a novel fuzzy-rough model and information granulation, Pattern Recognition 40 (2007) 3509–3521.
[14] Q.H. Hu, J.F. Liu, C.X. Wu, Neighborhood rough set based heterogeneous feature subset selection, Information Sciences 178 (18) (2008) 3577–3594.
[15] A. Kaufmann, Introduction to the Theory of Fuzzy Subsets, Academic Press, New York, 1975.
[16] G.J. Klir, Multivalued logics versus modal logics: alternative frameworks for uncertainty modelling, in: P.P. Wang (Ed.), Advances in Fuzzy Theory and Technology, Department of Electrical Engineering, Duke University, Durham, North Carolina, 1994, pp. 3–47.
[17] G.J. Klir, T.A. Fogler, Fuzzy Sets, Uncertainity and Information, Prentice-Hall Int. (UK) Ltd., 1998.

[18] B. Kosko, Neural Networks and Fuzzy Systems, Prentice-Hall, Englewood Cliffs, NJ, 1992.
[19] J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron, Rough sets: a tutorial, in: S.K. Pal, A. Skowron (Eds.), Rough Fuzzy Hybridization: A New Trend in Decision-Making, Springer-Verlag, Singapore, 1999, pp. 3–98.
[20] J.Y. Liang, K.S. Chin, C.Y. Dang, C.M. YAM. Richard, A new method for measuring uncertainty and fuzziness in rough set theory, International Journal of General Systems 31 (4) (2002) 331–342.
[21] J.Y. Liang, Z.Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (1) (2004) 37–46.
[22] J.Y. Liang, Z.Z. Shi, D.Y. Li, M.J. Wireman, The information entropy rough entropy and knowledge granulation in incomplete information systems, International Journal of General Systems 34 (1) (2006) 641–654.
[23] J.Y. Liang, J.H. Wang, Y.H. Qian, A new measure of uncertainty based on knowledge granulation for rough sets, Information Sciences 179 (4) (2009) 458–470.
[24] X.C. Liu, Entropy, distance measure and similarity measure of fuzzy sets and their relations, Fuzzy Sets and Systems 52 (1992) 305–318.
[25] A.D. Luca, S. Termini, A de1nition of a nonprobalistic entropy in the setting of fuzzy entropy, Information and Control 20 (1972) 301–312.
[26] D. Malyszko, J. Stepaniuk, Standard and fuzzy rough entropy clustering algorithms in image segmentation, Lecture Notes in Computer Science 5306 (2008) 409–418.
[27] D. Malyszko, J. Stepaniuk, Adaptive rough entropy clustering algorithms in image segmentation, Fundamenta Informaticae 98 (2–3) (2010) 199–231.
[28] D. Malyszko, J. Stepaniuk, Adaptive multilevel rough entropy evolutionary thresholding, Information Sciences 180 (2010) 1138–1158.
[29] J.S. Mi, W.X. Zhang, An aximatic characterization of a fuzzy generalization of rough sets, Information Sciences 160 (2004) 235–249.
[30] S. Nanda, Fuzzy rough sets, Fuzzy Sets and Systems 45 (1992) 157–160.
[31] N.R. Pal, S.K. Pal, Entropy, a new definition and its applications, IEEE Transactions on System, Man and Cybernetics 21 (1991) 1260–1270.
[32] N.R. Pal, J.C. Bezdek, Measuring fuzzy uncertainty, IEEE Transactions on Fuzzy Systems 2 (1994) 107–118.
[33] N.R. Pal, J.C. Bezdek, Quantifying different facets of fuzzy uncertainty, in: D. Dubas, H. Prade (Eds.), Fundamentals of Fuzzy Sets, Kluwer Academic Publishers., London, 2000, pp. 459–480.
[34] S.K. Pal, B.U. Shankar, P. Mitra, Granular computing rough entropy and object extraction, Pattern Recognition Letters 26 (2005) 2509–2517.
[35] Z. Pawlak, Information systems theoretical foundations, Information Systems 6 (3) (1981) 205–218.
[36] Z. Pawlak, Rough sets, International Journal of Computer and Information Sciences 11 (1982) 341–356.
[37] Z. Pawlak, Rough sets and fuzzy sets, Fuzzy Sets and Systems 17 (1985) 99–102.
[38] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Dordrecht, 1991.
[39] Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (2007) 3–27.
[40] Y.H. Qian, J.Y. Liang, C.Y. Dang, Consistency measure, inclusion degree and fuzzy measure in decision tables, Fuzzy Sets and Systems 159 (2008) 2353–2377.
[41] 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.
[42] Y.H. Qian, J.Y. Liang, W. Pedrycz, C.Y. Dang, Positive approximation: an accelerator for attribute reduction in rough set theory, Artificial Intelligence 174 (2010) 597–618.
[43] A.M. Radzikowska, E.E. Kerreb, A comparative study of fuzzy rough sets, Fuzzy Sets and Systems 126 (2002) 137–155.
[44] A. Renyi, On measures of entropy and information, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, vol. 1, 1960, pp. 547–561.
[45] C.E. Shannon, A mathematical theory of communication, Bell System Technology Journal 21 (1948) 379–423. 623–656.
[46] W. Wei, J.Y. Liang, Y.H. Qian, F. Wang, C.Y. Dang, Comparative study of decision performance of decision tables induced by attribute reductions, International Journal of General Systems 39 (8) (2010) 813–838.
[47] W.H. Xu, W.X. Zhang, Measuring roughness of generalized rough sets induced by a covering, Fuzzy Sets and Systems 158 (2007) 2443–2455.
[48] R.R. Yager, Measures of entropy and fuzziness related to aggregation operators, Information Sciences 82 (3–4) (1995) 147–166.
[49] R.R. Yager, On the measure of fuzziness and negation-Part II: Lattices, Information and Control 44 (1980) 236–260.
[50] R.R. Yager, On the measure of fuzziness and negation-Part I: Membership in the unit interval, International Journal of General Systems (5) (1979) 221–229.
[51] Y.J. Yang, R. John, Roughness bounds in rough set operations, Information Sciences 176 (2006) 3256–3267.
[52] J. Yao, M. Dash, S.T. Tan, H. Liu, Entropy-based fuzzy clustering and fuzzy modeling, Fuzzy Sets and Systems 113 (2000) 381–388.
[53] Y.Y. Yao, A comparative study of fuzzy sets and rough sets, Information Sciences 109 (1–4) (1998) 227–242.
[54] Y.Y. Yao, Constructive and algebraic methods of theory of rough sets, Information Sciences 109 (1998) 21–47.
[55] Y.Y. Yao, Neighborhood systems and approximate retrieval, Information Sciences 176 (23) (2006) 3431–3452.
[56] Y.Y. Yao, Attribute reduction in decision-theoretic rough set models, Information Sciences 178 (2008) 3356–3373.
[57] D.S. Yeung, D.G. Chen, E.C.C. Tsang, J.W.T. Lee, X.Z. Wang, On the generalization of fuzzy rough sets, IEEE Transaction on Fuzzy Systems 13 (3) (2005) 343–361.
[58] L.A. Zadeh, Fuzzy sets, Information Control 8 (1965) 338–353.
[59] L.A. Zadeh, Forward, in: E. Orlowska (Ed.), Incomplete Information: Rough Set Analysis, Physica-Verlag, Heidelberg, 1998, pp. v–vi.