Information Sciences 190 (2012) 1–16

Contents lists available at SciVerse ScienceDirect

Information Sciences

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

A comparative study of rough sets for hybrid data

Wei Wei, Jiye Liang *, Yuhua Qian

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

ARTICLE INFO

Article history:
Received 16 November 2010
Received in revised form 26 November 2011
Accepted 2 December 2011
Available online 9 December 2011

Keywords:
Fuzzy rough set
Neighborhood rough set
Hybrid data
Hybrid information granules
Granular computing

ABSTRACT

To discover knowledge from hybrid data using rough sets, researchers have developed several fuzzy rough set models and a neighborhood rough set model. These models have been applied to many hybrid data processing applications for a particular purpose, thus neglecting the issue of selecting an appropriate model. To address this issue, this paper mainly concerns the relationships among these rough set models. Investigating fuzzy and neighborhood hybrid granules reveals an important relationship between these two granules. Analyzing the relationships among rough approximations of these models shows that Hu's fuzzy rough approximations are special cases of neighborhood and Wang's fuzzy rough approximations, respectively. Furthermore, one-to-one correspondence relationships exist between Wang's fuzzy and neighborhood rough approximations. This study also finds that Wang's fuzzy and neighborhood rough approximations are cut sets of Dubois' fuzzy rough approximations and Radzikowska and Kerre's fuzzy rough approximations, respectively.

1. Introduction

In real world databases, data sets usually take on hybrid forms, i.e., the coexistence of categorical and numerical data. Feature selection, classification and prediction towards hybrid data thus hold great significance. Generally speaking, there are two strategies in hybrid data processing. One strategy is employing classical numerical data processing methods, including PCA [24], neural networks [6,14] and SVM [37]. When using these methods, all categorical data should be coded as integral numbers in hybrid data. However, processing categorical data in this manner is unreasonable, as the coded values of categorical data lack practical meanings [11]. Classical categorical data processing methods use the other strategy, including rough set theory [1,18,20–22,25,28,30–32,36,39,47]. Problems occur when numerical data are processed using traditional rough set theory. Discretizing numerical data into categorical data is thus necessary; however, this leads to the incurrence of information loss in the discretization process [11,46]. Both strategies mentioned above have their own limits.

Researchers have recently proposed several hybrid data processing methods [2,7,11,12,15,26,29,34,35,38,40], frequently using fuzzy and neighborhood rough set models. Fuzzy sets and rough sets are complementary in handling uncertainty [3,4,8,10,13,23,27,43]. Dubois and Prade [7] combined rough and fuzzy set theory to define the first fuzzy rough sets. This model employed the min and max fuzzy operators to describe the fuzzy lower and upper approximations. Radzikowska and Kerre [33] 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 triangular norm. Mi and Zhang [25] presented a new fuzzy rough set definition based on a residual implication θ and its dual σ. Hu et al. [11] introduced a novel fuzzy rough model, presented several attribute significance measures and designed a forward greedy algorithm for hybrid attribute reduction.

* 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).

0020-0255/$ - see front matter
doi:10.1016/j.ins.2011.12.006

Wang et al. [38] defined new lower and upper approximations based on the similarity between two objects and extended some underlying concepts to the fuzzy environment. Yeung et al. [46] first defined some lower and upper approximations based on arbitrary fuzzy relations from the constructive approach viewpoint. Some of the fuzzy rough set models mentioned above usually process hybrid data [7,11,35,38]. Furthermore, hybrid data analysis also employed another traditional rough set generalization: the neighborhood rough set [12,16,41,42,44,45]. Neighborhoods and neighborhood relations are important concepts in topology. Lin [19] regarded neighborhood spaces as general topological spaces more than equivalence spaces and introduced neighborhood relations into rough set methodology. The notion of neighborhood systems provided a convenient and flexible tool for representing similarity and described a hybrid information system with categorical and numerical attributes. Wu and Zhang [41] explicitly discussed the properties of neighborhood approximation spaces. Yao [43,45] relaxed the original query with a neighborhood system to conduct approximation retrieval. Hu et al. [12] constructed a unified theoretical framework for a neighborhood-based classifier using a neighborhood-based rough set model and a forward feature set selection algorithm towards hybrid data.

Some fuzzy and neighborhood rough set models mentioned above have been used to process hybrid data. However, a user cannot know which rough set model is appropriate when analyzing a given data set, making it difficult to select the appropriate model for a specific case. Solving this problem requires exploring the inherent relationships among the existing models, which helps researchers identify these generalized rough sets and select a proper model for a given application. This paper illustrates these relationships from two perspectives: constructing information granules and their rough approximations. It first discusses the analysis of the relationship between constructing fuzzy and neighborhood hybrid granules, in which information granules are the basis for rough approximations in rough set models. The paper then explores relationships among these rough approximations in the existing rough set models. This research clarifies the inherent relationships among these existing models.

The rest of the paper is organized as follows. Section 2 reviews some preliminary concepts. Section 3 analyzes the relationship between fuzzy hybrid granules and neighborhood hybrid granules. Section 4 introduces five rough set models for hybrid data. Section 5 investigates the relationships among the models, and the last section concludes the paper.

2. Preliminaries

Several fuzzy rough set models and the neighborhood rough set model are capable of processing hybrid data. To clarify the relationships among them, this section reviews some basic concepts, which facilitates the understanding of the remainder of this paper.

2.1. Hybrid information system

The hybrid information system occurs more frequently in real-world applications than does categorical information. A hybrid information system can be written as $(U, C^h = C^n \cup C^c)$, where U is the set of objects, C^n is a numerical attribute set and C^c is a categorical attribute set. To simplify this, we denote the *ith* numeric or categorical attribute in C^h as c_i^h. If every object in a hybrid information system belongs to a decision class generated from decision attribute D, the hybrid information system is a hybrid decision table, denoted as $(U, C^h \cup D)$.

2.2. Neighborhood rough set model

Let U be a finite universe. We associate each element $x \in U$ with a subset $n(x) \subseteq U$, called a neighborhood of x. A neighborhood of x may or may not contain x. A neighborhood system $NS(x)$ of x is a family of neighborhoods of x. A neighborhood operator $n : U \to 2^U$, where 2^U denotes the power set of the universe, can describe a neighborhood.

Let $n : U \to 2^U$ be a neighborhood operator. n is considered serial if, for all $x \in U$, there exists $y \in U$ such that $y \in n(x)$, i.e., for all $x \in U$, $n(x) \neq \emptyset$;n is considered inverse serial if, for all $x \in U$, there exists $y \in U$ such that $x \in n(y)$, i.e., $\cup_{x \in U} n(x) = U$; n is reflexive if, for all $x \in U$, $x \in n(x)$; n is symmetric if, for all $x,y \in U$, $x \in n(y)$ implies $y \in n(x)$; n is transitive if, for all $x,y,z \in U$, $y \in n(x)$ and $z \in n(y)$ imply $z \in n(x)$; and n is Euclidean if, for all $x,y,z \in U$, $y \in n(x)$ and $z \in n(x)$ imply $y \in n(z)$.

Combining these special properties, we can characterize various neighborhood systems [44]. In generalizing Pawlak's approximation operators, we use different neighborhood operators to define distinct approximation operators. For an equivalence relation R, the equivalence class $[x]_R$ may be considered a neighborhood of x. Let n denote an arbitrary neighborhood operator and $n(x)$ the corresponding neighborhood of x. Replacing $[X]_R$ with $n(x)$ in Pawlak's lower and upper approximations leads to the definition of a pair of approximation operators [44]:

$$\underline{apr}_n(X) = \{x | n(x) \subseteq X,\ x \in U\} \text{ and}$$
$$\overline{apr}_n(X) = \{x | n(x) \cap X \neq \emptyset,\ x \in U\},$$

where the subscript n indicates that the approximation operators are based on a particular neighborhood operator n. They can be viewed as a generalization of Pawlak's lower and upper approximations.

2.3. Fuzzy rough set model

Dubois and Prade first introduced the fuzzy rough set [7], hereafter called Dubois' fuzzy rough set for simplicity. According to their definition, a universe of objects $U = \{x_1, x_2, \ldots, x_n\}$ is described by a fuzzy binary relation $\widetilde{R}$, and the membership of object x_i in a fuzzy rough set $(\underline{\widetilde{R}}(A), \overline{\widetilde{R}}(A))$ is described as

$$\mu_{\underline{\widetilde{R}}(A)}(x_i) = \inf_{x_j \in U} \max\{1 - \widetilde{R}(x_i, x_j), \mu_A(x_j)\} \text{ and}$$
$$\mu_{\overline{\widetilde{R}}(A)}(x_i) = \sup_{x_j \in U} \min\{\widetilde{R}(x_i, x_j), \mu_A(x_j)\},$$

where $A \in \mathscr{F}(U)$. $\mathscr{F}(U)$ is the class of all fuzzy sets in U.

If $U/\widetilde{R} = \{F_1, F_2, \ldots, F_k\}$ is a fuzzy partition of U by a fuzzy binary relation $\widetilde{R}$, then the above expressions are equivalent to the following formulas [7]:

$$\mu_{\underline{\widetilde{R}}(A)}(F_i) = \inf_{x \in U} \max\{1 - \mu_{F_i}(x), \mu_A(x)\} \text{ and}$$
$$\mu_{\overline{\widetilde{R}}(A)}(F_i) = \sup_{x \in U} \min\{\mu_{F_i}(x), \mu_A(x)\}.$$

Furthermore, a collection of input fuzzy attributes $C_1, C_2, \ldots, C_m$, i.e., a set of fuzzy attributes, describes a universe of objects $U = \{x_1, x_2, \ldots, x_n\}$ [11]. Each fuzzy attribute contains a set of linguistic terms $F(C_i) = \{F_{ik} | k = 1, \ldots, P_{C_i}\}$, where P_{C_i} is the number of linguistic terms with respect to C_i. The set $U/C = \{F_{ik} | i = 1, \ldots, m; k = 1, \ldots, P_{C_i}\}$ can be regarded as fuzzy partitions of U by a set of fuzzy attributes C. For an arbitrary fuzzy set X, the membership degree of F_{ik} in the lower and upper approximations is

$$\mu_{\underline{\widetilde{R}}(A)}(F_{ik}) = \inf_{x \in U} \max\{1 - \mu_{F_{ik}}(x), \mu_A(x)\} \text{ and}$$
$$\mu_{\overline{\widetilde{R}}(A)}(F_{ik}) = \sup_{x \in U} \min\{\mu_{F_{ik}}(x), \mu_A(x)\}.$$

Radzikowska and Kerre presented a more general approach to the fuzzification of rough sets [33]. Furthermore, they introduced a broad family of fuzzy rough sets, each called an $(\mathscr{I}, \mathscr{T})$-fuzzy rough set, determined by an implicator $\mathscr{I}$ and triangular norm $\mathscr{T}$. The corresponding fuzzy approximation space and fuzzy rough approximations are defined below.

For a nonempty universe U and similarity relation $\widetilde{R}$ on U, a pair $S = (U, \widetilde{R})$ is called a fuzzy approximation space.

Let $S = (U, \widetilde{R})$ be a fuzzy approximation space and let $\mathscr{I}$ and $\mathscr{T}$ be a border implicator and a t-norm, respectively. The $(\mathscr{I}, \mathscr{T})$-fuzzy rough approximation in S is a mapping $Apr_S^{\mathscr{I},\mathscr{T}} : \mathscr{F}(U) \to \mathscr{F}(U) \times \mathscr{F}(U)$ defined by

for every $A \in \mathscr{F}(U)$

$$Apr_S^{\mathscr{I},\mathscr{T}} = ((\widetilde{R} \downarrow A)_{\mathscr{I}}(x), (\widetilde{R} \uparrow A)^{\mathscr{T}}(x))$$

and for every $x \in U$

$$(\widetilde{R} \downarrow A)_{\mathscr{I}}(x) = \inf_{y \in U} \mathscr{I}(\widetilde{R}(x, y), \mu_A(x)),$$
$$(\widetilde{R} \uparrow A)^{\mathscr{T}}(x) = \sup_{y \in U} \mathscr{T}(\widetilde{R}(x, y), \mu_A(x)),$$

where $\mathscr{F}(U)$ is the class of all fuzzy sets of U.

The implicator and t-norm notations are explained below.

A triangular norm, or t-norm, is an increasing, associative and commutative mapping $\mathscr{T} : [0,1]^2 \to [0,1]$ that satisfies the boundary condition ($\forall x \in [0,1]$, $T(x, 1) = x$). The most popular continuous t-norms are

- the standard min operator $\mathscr{T}_M(x, y) = \min\{x, y\}$,
- the algebraic product $\mathscr{T}_P(x, y) = x * y$,
- the bold intersection (also called the Łukasiewicz t-norm) $\mathscr{T}_L(x, y) = \max\{0, x + y - 1\}$.

A triangular conorm, or t-conorm, is an increasing, associative and commutative mapping $\mathscr{S} : [0,1]^2 \to [0,1]$ that satisfies the boundary condition ($\forall x \in [0,1]$, $\mathscr{S}(x, 0) = x$). Three well-known continuous conorms are

- the standard max operator $\mathscr{S}_M(x, y) = \max\{x, y\}$ (the smallest t-conorm),
- the probabilistic sum $\mathscr{S}_P(x, y) = x + y - x * y$,
- the bounded sum $\mathscr{S}_L(x; y) = \min\{1, x + y\}$.

A negator $\mathscr{N}$ is a decreasing $[0,1] - [0,1]$ mapping satisfying $\mathscr{N}(0) = 1$ and $\mathscr{N}(1) = 0$. The negator $\mathscr{N}_s = 1 - x$ is usually referred to as the standard negator. A negator $\mathscr{N}$ is involutive if $\mathscr{N}(\mathscr{N}(x)) = x$ for all $x \in [0,1]$, and it is weakly involutive if $\mathscr{N}(\mathscr{N}(x)) \geqslant x$ for all $x \in [0,1]$. Every involutive negator is continuous [17,33].

Let $\mathscr{T}, \mathscr{S}$ and $\mathscr{N}$ be a t-norm, t-conorm and negator, respectively. An implicator $\mathscr{I}$ is called an S-implicator based on $\mathscr{S}$ and $\mathscr{N}$ if $\mathscr{I}(x,y) = \mathscr{T}(\mathscr{N}(x), y)$ for all $x,y \in [0,1]$.

Three most popular S-implicators are

- the Łukasiewicz implicator $\mathscr{I}_L(x,y) = \min\{1, 1-x+y\}$, based on $\mathscr{S}_L$ and $\mathscr{N}_s$,
- the Kleene–Dienes implicator $\mathscr{I}_{KD}(x,y) = \max\{1-x, y\}$, based on $\mathscr{S}_M$ and $\mathscr{N}_s$,
- the Kleene–Dienes–Łukasiewicz implicator $\mathscr{I}_{\star}(x,y) = 1-x+x*y$, based on $\mathscr{S}_P$ and $\mathscr{N}_s$.

Hu presented another fuzzy rough set model [11], hereafter called Hu's fuzzy rough set, which specially processes hybrid data. The lower and upper approximations are based on the fuzzy hybrid granules, and they are given as follows.

Let $S = (U, C^h)$ be a hybrid information system and $X \subseteq U$ a crisp set of objects. The lower and upper approximations of X are

$$\underline{HC^h}(X) = \{x_i | [x_i]_{C^h} \subseteq X,\ x_i \in U\} \text{ and}$$

$$\overline{HC^h}(X) = \{x_i | [x_i]_{C^h} \cap X \neq \emptyset,\ x_i \in U\},$$

where $[x_i]_{C^h}$ is a hybrid granule with respect to C^h.

Wang proposed a new fuzzy rough set model [38], hereafter called Wang's fuzzy rough set, which is explicitly expressed thus:

Let $S = (U, C)$ be a fuzzy information system and $X \subseteq U$ a crisp subset of objects. Wang's fuzzy lower and upper approximations of X are

$$\underline{WC}_\beta(X) = \{x_i \in X | s_C(x_i, x_j) \leqslant 1-\beta, \forall x_j \in U - X\},$$

$$\overline{WC}_\beta(X) = \{x_i \in U | \exists x_j \in X,\ \text{such that } s_C(x_i, x_j) \geqslant \beta\},$$

where $s_C(x_i,x_j)$ is the similarity degree between x_i and x_j with respect to C.

As seen above, the neighborhood and fuzzy rough set models can process hybrid data. However, the inherent relationships among these existing models, which can help a researcher select a suitable model for a special case, have not yet been investigated. The following sections thus explore the relationships from two viewpoints: constructing information granules and their hybrid rough approximations.

3. Comparison of hybrid information granules

In this section, hybrid information granules are divided into two types: crisp and fuzzy hybrid granules. The following subsections explicitly investigate the relationship between them.

3.1. Construction of a crisp hybrid granule

The hybrid data can be divided into two parts: categorical and numerical. To construct the crisp hybrid granule, researchers introduced discretization algorithms to process the numerical part of the hybrid data. However, at least two structures are lost in the discretization process: the neighborhood and order structures in the numerical part. To solve the problem, Hu et al. introduced a neighborhood rough set model for hybrid attribute reduction [12].

Hybrid neighborhood granules can generally be constructed through the following three steps:

(1) constructing numerical neighborhood granules derived from a numerical attribute set;
(2) constructing categorical granules derived from a categorical attribute set;
(3) merging numerical and categorical neighborhood granules into hybrid neighborhood granules.

Hu et al. presented the following concrete method for constructing neighborhood granules [11,12].

Let $S = (U, C^h)$ be a hybrid information system, $C^n \subseteq C^h$ the numerical attribute set and $C^c \subseteq C^h$ the categorical attribute set; the numerical, categorical and hybrid neighborhood granules with respect to object $x \in U$ are defined as follows:

(1) $\delta_{C^n}(x_i) = \{x_j | d_{C^n}(x_i, x_j) \leqslant \delta,\ x_j \in U\}$,
(2) $\delta_{C^c}(x_i) = \{x_j | d_{C^c}(x_i, x_j) = 0,\ x_j \in U\}$,
(3) $\delta_{C^h}(x_i) = \{x_j | d_{C^n}(x_i, x_j) \leqslant \delta \wedge d_{C^c}(x_i, x_j) = 0,\ x_j \in U\}$.

where $d_{C^n}(x, x_i)$ is a distance function with respect to the numerical attribute subset, $d_{C^c}(x, x_i)$ is a distance function with respect to the categorical attribute subset and δ is a threshold.

3.2. Construction of a fuzzy hybrid granule

Whether objects are described in hybrid data by categorical or numeric attributes, a relation matrix can denote the relations between the objects. To construct a fuzzy hybrid granule, a fuzzy equivalence relation derived from each attribute $c_k^h \in C^h$ is introduced, and the relation matrix is indicated as follows.

$$M(R_{c_k^h}) = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{pmatrix},$$

where $r_{ij} \in [0,1]$ is the relation value of x_i and x_j with respect to c_k^h.

$R_{c_k^h}$ should satisfy:

(1) reflectivity: $R_{c_k^h}(x_i, x_i) = 1,\ \forall x_i \in X$;
(2) symmetry: $R_{c_k^h}(x_i, x_j) = R_{c_k^h}(x_j, x_i),\ \forall x_i, x_j \in X$;
(3) transitivity: $R_{c_k^h}(x_i, x_w) \geqslant \min_y\{R_{c_k^h}(x_i, x_j), R_{c_k^h}(x_j, x_w)\},\ \forall x_i, x_j, x_w \in X$.

The relation $R_{c_k^h}$ partitions U into many fuzzy hybrid granules (i.e., fuzzy equivalence classes) given by $U/R_{c_k^h} = \{[x_i]_{R_{c_k^h}}\}_{i=1}^n$ as $U/c_k^h = \{[x_i]_{c_k^h}\}_{i=1}^n$, where $[x_i]_{R_{c_k^h}}$ and $[x_i]_{c_k^h}$ denote the fuzzy equivalence classes determined by x_i with respect to a hybrid attribute c_k^h.

Hu et al. presented a concrete method for constructing fuzzy granules [11]. As with constructing crisp granules, generating numerical granules, creating categorical granules and merging numerical and categorical granules are necessary when constructing fuzzy hybrid granules, indicated as follows.

The fuzzy granule induced by a numerical attribute set is a fuzzy set in U, denoted as $[x_i]_{C^n} = \frac{r_{C^n}(x_i,x_1)}{x_1} + \frac{r_{C^n}(x_i,x_2)}{x_2} + \cdots + \frac{r_{C^n}(x_i,x_n)}{x_n}$, where $r_{C^n}(x_i, x_j) = \bigcap_{c_k^h \in C^n} r_{c_k^h}(x_i, x_j)$; the fuzzy granule derived from a categorical attribute set is a fuzzy set in U (in fact, it is a crisp set because the membership function belongs to {0,1}), denoted as $[x_i]_{C^c} = \frac{r_{C^c}(x_i,x_1)}{x_1} + \frac{r_{C^c}(x_i,x_2)}{x_2} + \cdots + \frac{r_{C^c}(x_i,x_n)}{x_n}$, where $r_{C^c}(x_i, x_j) = \bigcap_{c_k^h \in C^c} r_{c_k^h}(x_i, x_j)$; and a hybrid granule is generated by mixing numerical and categorical granules, denoted as $[x_i]_{C^h} = \frac{r_{C^h}(x_i,x_1)}{x_1} + \frac{r_{C^h}(x_i,x_2)}{x_2} + \cdots + \frac{r_{C^h}(x_i,x_n)}{x_n}$, where $r_{C^h}(x_i, x_j) = r_{C^n}(x_i, x_j) \cap r_{C^c}(x_i, x_j)$.

3.3. Comparing a neighborhood hybrid granule and a fuzzy hybrid granule

In a hybrid information system, the neighborhood hybrid granule with respect to a hybrid attribute set is a crisp object set, but the fuzzy hybrid granule (fuzzy equivalence class) induced by a hybrid attribute set is a fuzzy object set. These hybrid granules obviously differ. However, the above analyses (Sections 3.1 and 3.2) show that neighborhood and fuzzy hybrid granules are constructed based on the distance between objects. A relationship thus exists between a neighborhood hybrid granule and a fuzzy hybrid granule, given by the following theorem.

Theorem 3.1. *Let $S = (U, C^h)$ be a hybrid information system, $C^n \cup C^c = C^h$, C^n a numerical attribute set and C^c a categorical attribute set. If $r_{C^n}(x_i, x_j) = f(d_{C^n}(x_i, x_j)), r_{C^c}(x_i, x_j) = \begin{cases} 1, & d_{C^c}(x_i, x_j) = 0 \\ 0, & \text{otherwise} \end{cases}$ and $r_{C^h}(x_i, x_j) = \min\{r_{C^n}(x_i, x_j), r_{C^c}(x_i, x_j)\}$, then*

$$([x_i]_{C^h})_\lambda = \delta_{C^h}(x_i)|_{\delta = f^{-1}(\lambda)},$$

where $f(0) = 1,\ f(1) = 0, f(\cdot) \in [0,1],\ f(x) < f(y)$ if $x > y,\ f(x) = f(y)$ if $x = y,\ \lambda \in (0,1],\ \delta_{C^h}|_{\delta=f^{-1}(\lambda)}(x_i)$ indicates the neighborhood granule in which the parameter δ is equal to λ, and $d_{C^n}(x_i, x_j)$ and $d_{C^c}(x_i, x_j)$ are normalized distances, respectively.

Proof. According to the existing conditions, we have

$$r_{C^h}(x_i, x_j) = \min\{r_{C^n}(x_i, x_j), r_{C^c}(x_i, x_j)\} = \begin{cases} f(d_{C^n}(x_i, x_j)), & d_{C^c}(x_i, x_j) = 0, \\ 0, & \text{otherwise}. \end{cases}$$

Therefore,

$$\begin{aligned} ([x_i]_{C^h})_\lambda &= \{x_j | r_{C^h}(x_i, x_j) \geqslant \lambda, x_j \in U\} \\ &= \{x_j | (f(d_{C^n}(x_i, x_j)) \geqslant \lambda) \wedge (d_{C^c}(x_i, x_j) = 0), x_j \in U\} \\ &= \{x_j | d_{C^n}(x_i, x_j) \leqslant f^{-1}(\lambda) \wedge d_{C^c}(x_i, x_j) = 0, x_j \in U\} \\ &= \delta_{C^h}(x_i)|_{\delta = f^{-1}(\lambda)}. \qquad \square \end{aligned}$$

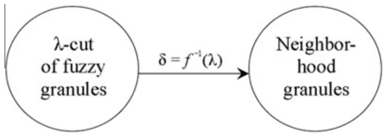

Fig. 1. Relationship between fuzzy granules and neighborhood granules.

Theorem 3.1 shows that the cut set of fuzzy hybrid granules is a neighborhood granule. Fig. 1 illustrates the relationship. Hu et al. presented a special case [11], in which the similarity between two objects is defined as

$$\begin{aligned} r_{C^h}(x_i,x_j) &= \min\{r_{C^n}(x_i,x_j), r_{C^c}(x_i,x_j)\} \\ &= \begin{cases} \min\{f(d_{C^n}(x_i,x_j)),1\}, & d_{C^n}(x_i,x_j) < \alpha \text{ and } d_{C^c}(x_i,x_j)=0 \\ 0, & \text{otherwise} \end{cases} \\ &= \begin{cases} 1-\frac{1}{\alpha}\times d_{C^n}(x_i,x_j), & d_{C^n}(x_i,x_j) < \alpha \text{ and } d_{C^c}(x_i,x_j)=0, \\ 0, & \text{otherwise}, \end{cases} \end{aligned}$$

where

$$r_{C^n}(x_i,x_j) = f(d_{C^n}(x_i,x_j)) = 1-\frac{1}{\alpha}\times d_{C^n}(x_i,x_j) \text{ and } r_{C^c}(x_i,x_j) = \begin{cases} 1, & d_{C^c}(x_i,x_j)=0 \\ 0, & \text{otherwise} \end{cases}.$$

According to Theorem 3.1, we get $([x_i]_{C^h})_\lambda = \delta_{C^h}(x_i)|_{(1-\lambda)\alpha}$. Furthermore, when the parameter $\alpha = 0.25$, the following equation can be obtained, $([x_i]_{C^h})_\lambda = \delta_{C^h}(x_i)|_{\delta=\frac{1-\lambda}{4}}$, where $\delta_{C^h}(x_i)|_{\delta=\frac{1-\lambda}{4}}$ indicates the neighborhood granule in which the threshold $\delta = \frac{1-\lambda}{4}$.

4. Rough approximations for hybrid data

Defining rough approximations (lower and upper approximations) is a key problem for a rough set model. In this section, we review several common rough approximations for hybrid data.

4.1. Neighborhood rough approximations

Hu et al. [11,12] applied the neighborhood rough set model to process hybrid information data, and the corresponding lower and upper approximations are defined thus:

Let $S = (U, C^h)$ be a hybrid information system and $X \subseteq U$ a crisp set of objects. The neighborhood lower and upper approximations of X can be defined as

$$\underline{NC^h}_\delta(X) = \{x_i | \delta_{C^h}(x_i) \subseteq X,\ x_i \in U\},$$
$$\overline{NC^h}_\delta(X) = \{x_i | \delta_{C^h}(x_i) \cap X \neq \emptyset,\ x_i \in U\}.$$

Furthermore, the lower and upper approximations of a hybrid decision table are thus:

Let $S = (U, C^h \cup D)$ be a hybrid decision table, C^h a hybrid condition attribute set, D a decision attribute, and $U/D = \{Y_1, Y_2, \ldots, Y_N\}$ a partition of discoursed universe U; the neighborhood lower and upper approximations for decision D are

$$\underline{NC^h}_\delta D = \cup_{i=1}^N \underline{NC^h}_\delta(Y_i),$$
$$\overline{NC^h}_\delta D = \cup_{i=1}^N \underline{NC^h}_\delta(Y_i).$$

4.2. Hu's fuzzy rough approximations

Hu's fuzzy rough set model is another rough set model for processing hybrid data. SubSection 2.3 introduced the definition of lower and upper approximations for the model.

Furthermore, for a given hybrid decision table $S = (U, C^h \cup D)$, $U/D = \{Y_1, Y_2, \ldots, Y_N\}$ is a partition of discoursed universe U. The lower and upper approximations with respect to the decision D are

$$\underline{HC^h}D = \cup_{i=1}^N \underline{HC^h}(Y_i),$$
$$\overline{HC^h}D = \cup_{i=1}^N \overline{HC^h}(Y_i).$$

4.3. Wang's fuzzy rough approximations

In hybrid information systems, the similarity degree between two objects with respect to a fuzzy attribute set C^h is $s(x_i, x_j)$, and Wang's fuzzy lower and upper approximations can be rewritten as

$$\underline{WC^h}_\beta(X) = \{x_i \in X | s(x_i, x_j) \leqslant 1 - \beta,\ \forall x_j \in U - X\},$$
$$\overline{WC^h}_\beta(X) = \{x_i \in U | \exists x_j \in X,\ \text{such that } s(x_i, x_j) \geqslant \beta\}.$$

For processing hybrid data using Wang's fuzzy rough set model, the similarity degree $r_{C^h}(x_i, x_j)$ is employed to measure the similarity between two objects. Thus, Wang's fuzzy lower and upper approximation can be rewritten as

$$\underline{WC^h}_\beta(X) = \{x_i \in X | r_{C^h}(x_i, x_j) \leqslant 1 - \beta, \forall x_j \in U - X\},$$
$$\overline{WC^h}_\beta(X) = \{x_i \in U | \exists x_j \in X,\ \text{such that } r_{C^h}(x_i, x_j) \geqslant \beta\}.$$

Furthermore, let $S = (U, C^h \cup D)$ be a hybrid decision table and $U/D = \{Y_1, Y_2, \ldots, Y_N\}$ a partition of discoursed universe U. The neighborhood lower and upper approximations for decision D are

$$\underline{WC^h}_\beta D = \cup_{i=1}^N \underline{WC^h}_\beta(Y_i),$$
$$\overline{WC^h}_\beta D = \cup_{i=1}^N \underline{WC^h}_\beta(Y_i).$$

The similarity between two objects in Wang's fuzzy lower and upper approximation can also apply to other similarity measures.

4.4. Dubois' fuzzy rough approximations

Hu et al. simplified Dubois' fuzzy rough approximations [9], where, for a given hybrid information system $S = (U, C^h)$, X is a crisp subset of U, and $r_{C^h}(x_i, x_j)$ measures the similarity between two objects, thus:

$$\mu_{\underline{C^h}(X)}(x_i) = \min_{x_j \notin X}\{1 - r_{C^h}(x_i, x_j)\},$$
$$\mu_{\overline{C^h}(X)}(x_i) = \max_{x_j \in X}\{r_{C^h}(x_i, x_j)\}.$$

4.5. Radzikowska and Kerre's fuzzy rough approximations

Cornelis et al. [5] used the model proposed by Radzikowska and Kerre [33] to obtain attribute reductions in hybrid data, using the Łukasievicz connectives $(\mathscr{T}_L, \mathscr{I}_L)$ and $(\mathscr{T}_M, \mathscr{I}_{KD})$. Let $S = (U, C^h)$ be a hybrid information system and X a crisp subset of U. $r_{C^h}(x_i, x_j)$ measures the similarity between two objects, rewriting Radzikowska and Kerre's fuzzy rough approximations as

$$(\widetilde{R} \downarrow X)_{\mathscr{I}_L}(x_i) = \inf_{x_j \in U} \mathscr{I}_L(r_{C^h}(x_i, x_j), \mu_X(x_i)),$$
$$(\widetilde{R} \uparrow X)^{\mathscr{T}_L}(x_i) = \sup_{x_j \in U} \mathscr{T}_L(r_{C^h}(x_i, x_j), \mu_X(x_i)),$$
$$(\widetilde{R} \downarrow X)_{\mathscr{I}_{KD}}(x_i) = \inf_{x_j \in U} \mathscr{I}_{KD}(r_{C^h}(x_i, x_j), \mu_X(x_i)),$$
$$(\widetilde{R} \uparrow X)^{\mathscr{T}_M}(x_i) = \sup_{x_j \in U} \mathscr{T}_M(r_{C^h}(x_i, x_j), \mu_X(x_i)).$$

5. Comparing rough approximations for hybrid data

Hu's fuzzy, neighborhood and Wang's fuzzy rough approximations for hybrid data are all crisp object sets, whereas Dubois' and Radzikowska and Kerre's fuzzy rough approximations for hybrid data are fuzzy object sets. These rough approximations are divided into two types: crisp and fuzzy hybrid rough approximations. This section investigates the relationships among them.

5.1. Relationships among crisp hybrid rough approximations

Neighborhood rough approximations are defined based on neighborhood hybrid granules, and Hu's fuzzy rough approximations are defined by constructing fuzzy hybrid granules. Furthermore, because the cut sets of a fuzzy hybrid granule is a neighborhood hybrid granule, neighborhood rough approximations are more general than Hu's fuzzy ones. The following theorem offers a concrete explanation.

Theorem 5.1. *Let $S = (U, C^h)$ be a hybrid information system and let $X \subseteq U$ be a crisp set. If $r_{C^n}(x_i, x_j) = f(d_{C^n}(x_i, x_j)), r_{C^c}(x_i, x_j) = \begin{cases} 1, & d_{C^c}(x_i, x_j) = 0 \\ 0, & \text{otherwise} \end{cases}$ and $r_{C^h}(x_i, x_j) = \min\{r_{C^n}(x_i, x_j), r_{C^c}(x_i, x_j)\}$, then*

$$\underline{NC^h}_\xi(X) = \underline{HC^h}(X) \text{ and } \overline{NC^h}_\xi(X) = \overline{HC^h}(X),$$

where ξ *is a constant that satisfies* $\delta_{C^h}(x_i)|_{\delta=\xi} = \delta_{C^h}(x_i)|_{\delta=1} - \{x_j | d_{C^h}(x_i,x_j) = 1\}$.

Proof. From Theorem 3.1 and the existing conditions, we have

$$\delta_{C^h}(x_i)|_{\delta=1} = ([x_i]_{C^h})_0.$$

Furthermore, because $\{x_j | d_{C^h}(x_i,x_j) = 1\} = \{x_j | r_{C^h}(x_i,x_j) = 0\}$,

$$\delta_{C^h}(x_i)|_{\delta=\xi} = \delta_{C^h}(x_i)|_{\delta=1} - \{x_j | d_{C^h}(x_i,x_j) = 1\} = ([x_i]_{C^h})_0 - \{x_j | r_{C^h}(x_i,x_j) = 0\}.$$

Therefore,

$$\begin{aligned}\underline{NC^h}_\xi(X) &= \{x_i | \delta_{C^h}(x_i)|_{\delta=\xi} \subseteq X, x_i \in U\} \\ &= \left\{x_i \middle| \left(([x_i]_{C^h})_0 - \{x_j | r_{C^h}(x_i,x_j) = 0\}\right) \subseteq X, x_i \in U\right\} \\ &= \{x_i | [x_i]_{C^h} \subseteq X, x_i \in U\} \\ &= \underline{HC^h}(X) \text{ and}\end{aligned}$$

$$\begin{aligned}\overline{NC^h}_\xi(X) &= \{x_i | \delta_{C^h}(x_i)|_{\delta=\xi} \cap X \neq \emptyset, x_i \in U\} \\ &= \left\{x_i \middle| \left(([x_i]_{C^h})_0 - \{x_j | r_{C^h}(x_i,x_j) = 0\}\right) \cap X \neq \emptyset, x_i \in U\right\} \\ &= \{x_i | [x_i]_{C^h} \cap X \neq \emptyset, x_i \in U\} \\ &= \overline{HC^h}(X).\end{aligned}$$

Theorem 5.1 shows that, in essence, neighborhood rough approximations are identical to Hu's if the parameter satisfies a special condition. Neighborhood lower and upper approximations are the generalizations of Hu's fuzzy ones. It can be specifically indicated by Fig. 2.

Furthermore, the definitions indicate that Wang's fuzzy rough approximations are generated using the similarity between two objects, and fuzzy hybrid granules construct Hu's fuzzy rough approximations, which also rely on the similarity between two objects. Some relationships among Wang's and Hu's fuzzy rough approximations may therefore exist. The following two theorems investigate these relationships.

Theorem 5.2. *Let S = (U, C^h) be a hybrid information system, and* $X \subseteq U$*. If* $r_{C^h}(x_i,x_j)$ *measures the similarity between two objects. If*

$$r_{C^n}(x_i,x_j) = f(d_{C^n}(x_i,x_j)), r_{C^c}(x_i,x_j) = \begin{cases} 1, & d_{C^c}(x_i,x_j) = 0 \\ 0, & \text{otherwise} \end{cases} \text{ and } r_{C^h}(x_i,x_j) = \min\{r_{C^n}(x_i,x_j), r_{C^c}(x_i,x_j)\},$$

then

$$\underline{WC^h}_1(X) = \underline{HC^h}(X).$$

Proof. From the definition of Wang's lower and upper approximations, we have

$$\begin{aligned}\underline{WC^h}_1(X) &= \{x_i \in X | r_{C^h}(x_i,x_j) \leqslant 1 - 1, \ \forall x_j \notin X\} \\ &= \{x_i \in X | r_{C^h}(x_i,x_j) = 0, \ \forall x_j \notin X\} \\ &= \{x_i \in X | x_j \in X \text{ if } r_{C^h}(x_i,x_j) > 0\} \\ &= \{x_i | [x_i]_{C^h} \subseteq X, x_i \in U\} \\ &= \underline{HC^h}(X). \qquad \square\end{aligned}$$

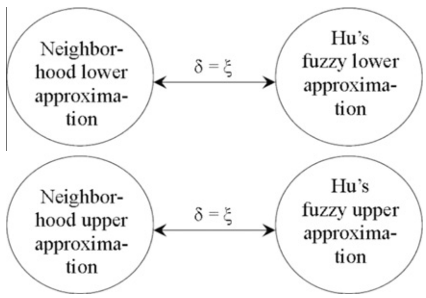

Fig. 2. Relationships between neighborhood rough approximations and Hu's fuzzy rough approximations.

Theorem 5.2 shows that Wang's fuzzy lower approximation is the same as Hu's if the parameter $\beta = 1$ in Wang's lower approximation. Therefore, we conclude that Wang's fuzzy lower approximation is more general than Hu's.

Theorem 5.3. *Let $S = (U, C^h)$ be a hybrid information system, and $X \subseteq U$. If $r_{C^h}(x_i, x_j)$ evaluates the similarity between two objects in Wang's and Hu's fuzzy upper approximations, then*

$$\overline{WC^h}_\zeta(X) = \overline{HC^h}(X),$$

where ζ is a constant that satisfies $\{x_i \in U | r_{C^h}(x_i, x_j) \geqslant \zeta,\ \exists x_j \in X\} = \{x_i \in U | r_{C^h}(x_i, x_j) \geqslant 0,\ \exists x_j \in X\} - \{x_i \in U | r_{C^h}(x_i, x_j) = 0,\ \exists x_j \in X\}$.

Proof. From the existing condition, we have

$$\begin{aligned}\overline{WC^h}_\zeta(X) &= \{x_i \in U | r_{C^h}(x_i, x_j) \geqslant \zeta, \exists x_j \in X\} \\ &= \{x_i | r_{C^h}(x_i, x_j) > 0, \exists x_j \in X\} \\ &= \{x_i | [x_i]_{C^h} \cap X \neq \emptyset, \exists x_j \in X\} \\ &= \overline{HC^h}(X). \qquad \square\end{aligned}$$

Theorem 5.3 indicates that Hu's fuzzy upper approximation is identical to Wang's if the parameter β is an infinitesimal. Therefore, in some sense, Hu's fuzzy upper approximation is a special case of Wang's. Fig. 3 specifically indicates the results from Theorems 5.2 and 5.3.

The following two theorems examine the relationships among Wang's fuzzy rough approximations and neighborhood rough approximations.

Theorem 5.4. *Let $S = (U, C^h)$ be a hybrid information system and $X \subseteq U$ a crisp set. If $r_{C^n}(x_i, x_j) = f(d_{C^n}(x_i, x_j)), r_{C^c}(x_i, x_j) = \begin{cases} 1, & d_{C^c}(x_i, x_j) = 0 \\ 0, & \text{otherwise} \end{cases}$, and $r_{C^h}(x_i, x_j) = \min\{r_{C^n}(x_i, x_j), r_{C^c}(x_i, x_j)\}$, then*

$$\underline{WC^h}_\beta(X) = \underline{NC^h}_{f^{-1}(1-\beta)}(X) - \{x_i \in X | x_j \in X \text{ if } (d_{C^h}(x_i, x_j) = f^{-1}(1-\beta))\},$$

where $f(0) = 1$, $f(1) = 0$, $f(\cdot) \in [0,1]$, $f(x) < f(y)$ if $x > y$, and $f(x) = f(y)$ if $x = y$.

Proof. From the existing conditions, we have

$$\begin{aligned}\underline{WC^h}_\beta(X) &= \{x_i \in X | r_{C^h}(x_i, x_j) \leqslant 1 - \beta, \forall x_j \notin X\} \\ &= \{x_i \in X | d_{C^n}(x_i, x_j) \geqslant f^{-1}(1-\beta) \wedge d_{C^c}(x_i, x_j) = 0,\ \forall x_j \notin X\} \\ &= \{x_i \in X | x_j \in X \text{ if } d_{C^n}(x_i, x_j) < f^{-1}(1-\beta) \wedge d_{C^c}(x_i, x_j) = 0\} \\ &= \{x_i \in X | x_j \in X \text{ if}(d_{C^n}(x_i, x_j) \leqslant f^{-1}(1-\beta)) \wedge d_{C^c}(x_i, x_j) = 0\} \\ &\quad - \{x_i \in X | x_j \in X \text{ if}(d_{C^n}(x_i, x_j) = f^{-1}(1-\beta)) \wedge d_{C^c}(x_i, x_j) = 0\} \\ &= \{x_i | \delta_{C^h}(x_i)|_{f^{-1}(1-\beta)} \subseteq X, x_i \in U\} \\ &\quad - \{x_i \in X | x_j \in X \text{ if } (d_{C^n}(x_i, x_j) = f^{-1}(1-\beta)) \wedge d_{C^c}(x_i, x_j) = 0\} \\ &= \underline{NC^h}_{f^{-1}(1-\beta)}(X) - \{x_i \in X | x_j \in X \text{ if } (d_{C^h}(x_i, x_j) = f^{-1}(1-\beta))\}. \qquad \square\end{aligned}$$

Theorem 5.4 indicates a one to one correspondence between Wang's fuzzy lower approximation and neighborhood lower approximation.

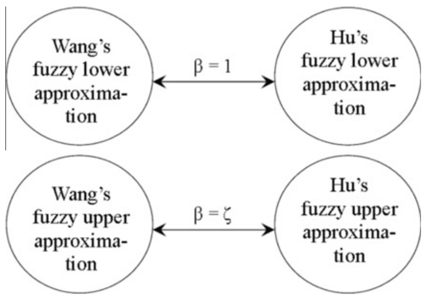

Fig. 3. Relationships between Wang's rough approximations and Hu's fuzzy rough approximations.

Theorem 5.5. *Let $S = (U, C^h)$ be a hybrid information system and let $X \subseteq U$ be a crisp set, and $r_{C^h}(x_i, x_j)$ measures the similarity between two objects. If $r_{C^n}(x_i, x_j) = f(d_{C^n}(x_i, x_j)), r_{C^c}(x_i, x_j) = \begin{cases} 1, & d_{C^c}(x_i, x_j) = 0 \\ 0, & otherwise \end{cases}$ and $r_{C^h}(x_i, x_j) = \min\{r_{C^n}(x_i, x_j), r_{C^c}(x_i, x_j)\}$, then*

$$\overline{WC^h}_\beta(X) = \overline{NC^h}_{f^{-1}(\beta)}(X),$$

where $f(0) = 1$, $f(1) = 0$, $f(\cdot) \in [0,1]$, $f(x) < f(y)$ if $x > y$, and $f(x) = f(y)$ if $x = y$.

Proof. According to the existing condition, we have

$$\begin{aligned} \overline{WC^h}_\beta(X) &= \{x_i \in U | r_{C^h}(x_i, x_j) \geqslant \beta, \exists x_j \in X\} \\ &= \{x_i | d_{C^h}(x_i, x_j) \leqslant f^{-1}(\beta), \exists x_j \in X\} \\ &= \{x_i | \delta_{C^h}(x_i)|_{\delta = f^{-1}(\beta)} \cap X \neq \emptyset\} \\ &= \overline{NC^h}_{f^{-1}(\beta)}(X). \quad \square \end{aligned}$$

As in Theorems 5.4 and 5.5 suggests that there is a one to one correspondence between Wang's fuzzy upper approximation and neighborhood upper approximation. Fig. 4 illustrates these relationships.

The conclusions in Theorems 5.4 and 5.5 indicate that the one-to-one correspondence between two lower approximations differ from that between the two upper approximations. The following theorems give the reason for this problem.

Theorem 5.6. *Let $S = (U, C^h)$ be a hybrid information system and let $X \subseteq U$ be a crisp set. If $\beta_1 > \beta_2$, then*

$$\underline{WC^h}_{\beta_1}(X) \subseteq \underline{WC^h}_{\beta_2}(X),$$
$$\overline{WC^h}_{\beta_1}(X) \subseteq \overline{WC^h}_{\beta_2}(X).$$

Proof. From the existing conditions, we have

$$\begin{aligned} \underline{WC^h}_{\beta_1}(X) &= \{x_i \in X | r_{C^h}(x_i, x_j) \leqslant 1 - \beta_1, \forall x_j \notin X\} \\ &\subseteq \{x_i \in X | r_{C^h}(x_i, x_j) \leqslant 1 - \beta_2, \forall x_j \notin X\} \\ &= \{x_i \in X | x_j \in X \text{ if } r_{C^h}(x_i, x_j) > 1 - \beta_2\} \\ &= \underline{WC^h}_{\beta_2}(X) \text{ and} \end{aligned}$$

$$\begin{aligned} \overline{WC^h}_{\beta_1}(X) &= \{x_i \in U | r_{C^h}(x_i, x_j) \geqslant \beta_1, \exists x_j \in X\} \\ &= \{x_i \in U | r_{C^h}(x_i, x_j) > \beta_2, \exists x_j \in X\} \\ &\subseteq \{x_i \in U | r_{C^h}(x_i, x_j) \geqslant \beta_2, \exists x_j \in X\} \\ &= \overline{WC^h}_{\beta_2}(X). \end{aligned}$$

Theorem 5.7. *Let $S = (U, C^h)$ be a hybrid information system and let $X \subseteq U$ be a crisp set. If $\delta_1 > \delta_2$, then*

$$\underline{NC^h}_{\delta_1}(X) \subseteq \underline{NC^h}_{\delta_2}(X),$$
$$\overline{NC^h}_{\delta_1}(X) \supseteq \overline{NC^h}_{\delta_2}(X).$$

As in the proof of Theorem 5.6, using the definitions of neighborhood lower and upper approximations, proving the theorem is easy.

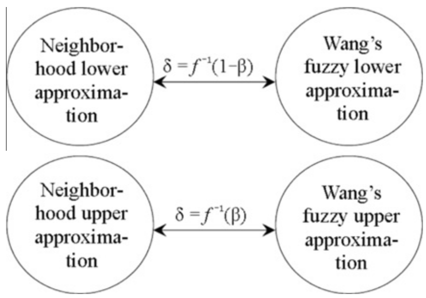

Fig. 4. Relationships between Wang's fuzzy rough approximations and neighborhood rough approximations.

Theorems 5.6 and 5.7 show that changing Wang's fuzzy rough approximations with parameter β differs from changing neighborhood rough approximations with parameter δ.

5.2. Relationships among fuzzy and crisp hybrid rough approximations

Both fuzzy and crisp hybrid rough approximations are constructed based on the similarity between two objects. Therefore, we speculate that there exists some inherent relationships among them. Dubois' and Radzikowska and Kerre's fuzzy rough approximations are two important fuzzy hybrid rough approximations. This subsection therefore investigates the relationships among crisp and fuzzy hybrid rough approximations.

In the following, several theorems illustrate the relationships among Dubois' fuzzy rough approximations and crisp hybrid rough approximations.

Theorem 5.8. *Let $S=(U,C^h)$ be a hybrid information system, and $X\subseteq U$. If $r_{C^n}(x_i,x_j)=f(d_{C^n}(x_i,x_j)), r_{C^c}(x_i,x_j)=\begin{cases}1, & d_{C^c}(x_i,x_j)=0\\ 0, & \text{otherwise}\end{cases}$ and $r_{C^h}(x_i,x_j)=\min\{r_{C^n}(x_i,x_j),r_{C^c}(x_i,x_j)\}$, then*

$$(\underline{C^h}(X))_1=\underline{HC^h}(X),$$
$$(\overline{C^h}(X))_0^S=\overline{HC^h}(X).$$

Proof. According to the existing conditions, we have

$$\begin{aligned}(\underline{C^h}(X))_1&=\{x_i|\min_{x_j\notin X}\{1-r_{C^h}(x_i,x_j)\}\geqslant 1\}\\&=\{x_i|1-r_{C^h}(x_i,x_j)\geqslant 1,\forall x_j\notin X\}\\&=\{x_i|r_{C^h}(x_i,x_j)\leqslant 0,\forall x_j\notin X\}\\&=\{x_i|r_{C^h}(x_i,x_j)=0,\forall x_j\notin X\}\\&=\{x_i|[x_i]_{C^h}\subseteq X,x_i\in U\}\\&=\underline{HC^h}(X),\end{aligned}$$

$$\begin{aligned}(\overline{C^h}(X))_0^S&=\{x_i|\max_{x_j\in X}\{r_{C^h}(x_i,x_j)\}>0\}\\&=\{x_i|r_{C^h}(x_i,x_j)>0,\exists x_j\in X\}\\&=\{x_i|[x_i]_{C^h}\cap X\neq\emptyset,x_i\in U\}\\&=\overline{HC^h}(X).\end{aligned}$$

Theorem 5.8 shows that Hu's lower approximation is the 1-cut of Dubois' fuzzy lower approximation, and Hu's upper approximation is the strong 0-cut of Dubois' fuzzy upper approximation.

Theorem 5.9. *Let $S=(U,C^h)$ be a hybrid information system, and $X\subseteq U$. If $r_{C^n}(x_i,x_j)=f(d_{C^n}(x_i,x_j)), r_{C^c}(x_i,x_j)=\begin{cases}1, & d_{C^c}(x_i,x_j)=0\\ 0, & \text{otherwise}\end{cases}$ and $r_{C^h}(x_i,x_j)=\min\{r_{C^n}(x_i,x_j),r_{C^c}(x_i,x_j)\}$, then*

$$(\underline{C^h}(X))_\lambda=\underline{NC^h}_{f^{-1}(1-\lambda)}(X)-\{x_i\in X|x_j\in X\ \mathit{if}(d_{C^n}(x_i,x_j)=f^{-1}(1-\lambda))\wedge d_{C^c}(x_i,x_j)=0\},$$

where $\lambda\in(0,1]$, $\delta\in[0,1)$, $f(0)=1$, $f(1)=0$, $f(\cdot)\in[0,1]$, $f(x)<f(y)$ if $x>y$, and $f(x)=f(y)$ if $x=y$.

Proof. According to the existing conditions, we have

$$\begin{aligned}(\underline{C^h}(X))_\lambda&=\{x_i|\min_{x_j\notin X}\{1-r_{C^h}(x_i,x_j)\}\geqslant\lambda\}\\&=\{x_i|1-r_{C^h}(x_i,x_j)\geqslant\lambda,\forall x_j\notin X\}\\&=\{x_i|r_{C^h}(x_i,x_j)\leqslant 1-\lambda,\forall x_j\notin X\}\\&=\{x_i|(d_{C^n}(x_i,x_j)\geqslant f^{-1}(1-\lambda)\wedge d_{C^c}(x_i,x_j)=0)\vee(d_{C^c}(x_i,x_j)=1),\forall x_j\notin X\}\\&=\{x_i\in X|x_j\in X\ \text{if}(d_{C^n}(x_i,x_j)\leqslant f^{-1}(1-\lambda))\wedge d_{C^c}(x_i,x_j)=0\}\\&\quad-\{x_i\in X|x_j\in X\ \text{if}(d_{C^n}(x_i,x_j)=f^{-1}(1-\lambda))\wedge d_{C^c}(x_i,x_j)=0\}\\&=\{x_i|\delta_{C^h}(x_i)|_{\delta=f^{-1}(1-\lambda)}\subseteq X,x_i\in U\}\\&\quad-\{x_i\in X|x_j\in X\ \text{if}(d_{C^n}(x_i,x_j)=f^{-1}(1-\lambda))\wedge d_{C^c}(x_i,x_j)=0\}\\&=\underline{NC^h}_{f^{-1}(1-\lambda)}(X)-\{x_i\in X|x_j\in X\ \text{if}(d_{C^n}(x_i,x_j)=f^{-1}(1-\lambda))\wedge d_{C^c}(x_i,x_j)=0\}.\end{aligned}$$

Theorem 5.9 indicates that neighborhood lower approximation is inherently identical to the λ-cut of Dubois' fuzzy lower approximation because a one-to-one correspondence between parameters δ and λ.

Theorem 5.10. *Let $S=(U,C^h)$ be a hybrid information system, and $X\subseteq U$. If $r_{C^n}(x_i,x_j)=f(d_{C^n}(x_i,x_j)), r_{C^c}(x_i,x_j)=\begin{cases}1, & d_{C^c}(x_i,x_j)=0\\ 0, & otherwise\end{cases}$, and $r_{C^h}(x_i,x_j)=\min\{r_{C^n}(x_i,x_j),r_{C^c}(x_i,x_j)\}$, then*

$$(\overline{C^h}(X))_\lambda=\overline{NC^h}_{f^{-1}(\lambda)}(X),$$

where $\lambda\in(0,1]$, $\delta\in[0,1)$, $f(0)=1$, $f(1)=0$, $f(\cdot)\in[0,1]$, $f(x)<f(y)$ if $x>y$, and $f(x)=f(y)$ if $x=y$.

Proof. According to the existing condition, we have

$$\begin{aligned}(\overline{C^h}(X))_\lambda &=\{x_i|\max_{x_j\in X}\{r_{C^h}(x_i,x_j)\}\geqslant\lambda\}\\ &=\{x_i|r_{C^h}(x_i,x_j)\geqslant\lambda,\exists x_j\in X\}\\ &=\{x_i|d_{C^n}(x_i,x_j)\leqslant f^{-1}(\lambda)\wedge d_{C^c}(x_i,x_j)=0,\exists x_j\in X\}\\ &=\overline{NC^h}_{f^{-1}(\lambda)}(X).\qquad\square\end{aligned}$$

Similar to Theorems 5.9, 5.10 shows that neighborhood upper approximation is the same as the λ-cut of Dubois' fuzzy upper approximation because a one-to-one correspondence also exists between δ and λ.

Theorem 5.11. *Let $S=(U,C^h)$ be a hybrid information system, and $X\subseteq U$. If $r_{C^h}(x_i,x_j)$ evaluates the similarity between two objects in Dubois' and Wang's fuzzy upper approximations, then*

$$(\underline{C^h}(X))_\lambda=\underline{WC^h}_\lambda(X),$$
$$(\overline{C^h}(X))_\lambda=\overline{WC^h}_\lambda(X).$$

Proof. From the existing condition, we have

$$\begin{aligned}(\underline{C^h}(X))_\lambda &=\{x_i|\min_{x_j\notin X}\{1-r_{C^h}(x_i,x_j)\}\geqslant\lambda\}\\ &=\{x_i|1-r_{C^h}(x_i,x_j)\geqslant\lambda,\forall x_j\notin X\}\\ &=\{x_i|r_{C^h}(x_i,x_j)\leqslant 1-\lambda,\forall x_j\notin X\}\\ &=\underline{WC^h}_\lambda(X).\end{aligned}$$

Furthermore,

$$\begin{aligned}(\overline{C^h}(X))_\lambda &=\{x_i|\max_{x_j\in X}\{r_{C^h}(x_i,x_j)\}\geqslant\lambda\}\\ &=\{x_i|r_{C^h}(x_i,x_j)\geqslant\lambda,\exists x_j\in X\}\\ &=\overline{WC^h}_\lambda(X).\qquad\square\end{aligned}$$

Theorem 5.11 states that Wang's fuzzy rough approximations is in essence equal to the λ-cut of Dubois' fuzzy rough approximations.

Fig. 5 illustrates the relationships among Hu's fuzzy, neighborhood, Wang's fuzzy and Dubois' fuzzy rough approximations.

Radzikowska and Kerre demonstrated that employing $(\mathscr{T}_M,\mathscr{I}_{KD})$ in Radzikowska and Kerre's fuzzy rough approximations gives exactly Dubois' fuzzy rough approximations [33]. Fig. 6 illustrates this relationship. Furthermore, using the results from Theorems 5.8, 5.9, 5.10, 5.11, obtaining relationships among Radzikowska and Kerre's fuzzy and crisp hybrid rough approximations is easy.

Example 1 better illustrates the relationships among crisp and fuzzy hybrid rough approximations. The above analyses show that the relationships among Dubois' fuzzy and crisp hybrid approximations are representative. Therefore, we only analyze the relationships among Dubois' fuzzy and crisp hybrid rough approximations in the example.

Example 1. Table 1 is part of the table Ecoli in UCI datasets, in which *Sequence name* is the ID of objects, *MCG*, *GVH*, *LIP*, *CHG*, *AAC*, *ALM*1 and *ALM*2 are the condition attributes (*LIP* and *CHG* are categorical, and the others are numerical), and *Class* is the decision attribute. Table 1 indicates that it is a hybrid decision table. For convenience, suppose that $C^n=\{MCG,GVH,AAC,ALM1,ALM2\}$, $C^c=\{LIP,CHG\}$, $C^h=C^n\cup C^c$ and $D=\{Class\}$.

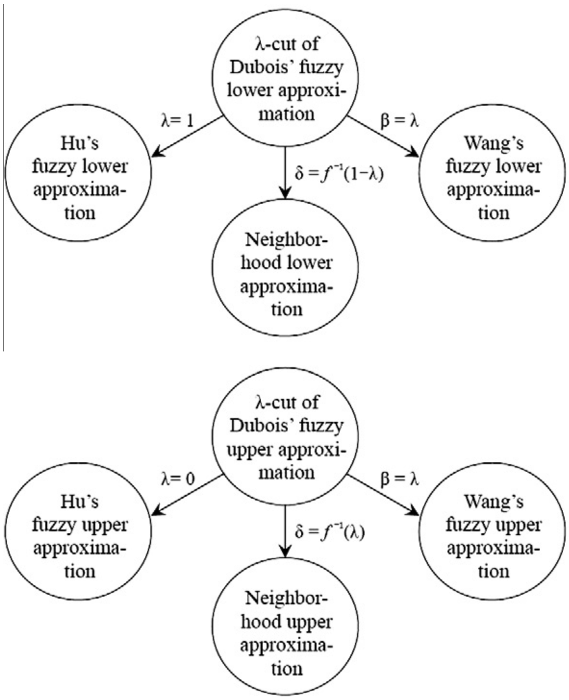

Fig. 5. Relationships among Dubois' fuzzy rough approximation and crisp hybrid rough approximations.

Without loss of generality, let $r_{C^n}(x_i,x_j)=f(d_{C^n}(x_i,x_j))=\begin{cases}1-2d_{C^n}(x_i,x_j), & d_{C^n}(x_i,x_j)<0.5\\ 0, & \text{otherwise}\end{cases}$ and $r_{C^c}(x_i,x_j)=\begin{cases}1, & d_{C^c}(x_i,x_j)=0\\ 0, & \text{otherwise}\end{cases}$, where $d_{C^n}(x_i,x_j)=\max_{k=1}^{|C^n|}\left\{\frac{|f(x_i,c_k^n)-f(x_j,c_k^n)|}{\max_{x_l\in U}\{f(x_l,c_k^n)\}}\right\}$ and $d_{C^c}(x_i,x_j)=\begin{cases}0, & f(x_i,c_k^c)=f(x_j,c_k^c) for \forall c_k^c\in C^c\\ 1, & \text{otherwise}\end{cases}$.

After computing, we obtain the following distance matrix:

$$D(C^h)=\begin{pmatrix}
0 & 0.2323 & 0.3529 & 0.7176 & 0.7412 & 0.8353 & 1 & 0.7412 & 1 & 1\\
0.2323 & 0 & 0.2353 & 0.6406 & 0.6235 & 0.7879 & 0.8824 & 0.6235 & 1 & 1\\
0.3529 & 0.2353 & 0 & 0.6563 & 0.3882 & 0.6566 & 0.6471 & 0.4545 & 1 & 1\\
0.7176 & 0.6406 & 0.6563 & 0 & 1 & 1 & 0.8281 & 0.8438 & 1 & 1\\
0.7412 & 0.6235 & 0.3882 & 1 & 0 & 0.3334 & 0.3131 & 0.3467 & 1 & 1\\
0.8353 & 0.7879 & 0.6566 & 1 & 0.3334 & 0 & 0.6465 & 0.2400 & 1 & 1\\
1 & 0.8824 & 0.6470 & 0.8281 & 0.3131 & 0.6465 & 0 & 0.4444 & 1 & 1\\
0.7412 & 0.6235 & 0.4545 & 0.8438 & 0.3467 & 0.2400 & 0.4444 & 0 & 1 & 1\\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1\\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0
\end{pmatrix},$$

and similarity matrix:

$$M(C^h)=\begin{pmatrix}
1 & 0.5354 & 0.2941 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0.5354 & 1 & 0.5294 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0.2941 & 0.5294 & 1 & 0 & 0.2235 & 0 & 0 & 0.0909 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0.2235 & 0 & 1 & 0.3334 & 0.3737 & 0.3067 & 0 & 0\\
0 & 0 & 0 & 0 & 0.3334 & 1 & 0 & 0.5200 & 0 & 0\\
0 & 0 & 0 & 0 & 0.3737 & 0 & 1 & 0.1111 & 0 & 0\\
0 & 0 & 0.0909 & 0 & 0.3067 & 0.5200 & 0.1111 & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.$$

The equivalent classes induced by hybrid attribute set C^h and decision attribute *Class* are

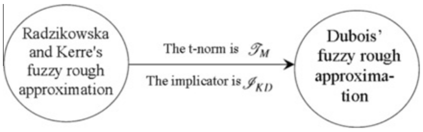

Fig. 6. Relationship between Dubois' and Radzikowska and Kerre's fuzzy rough approximations.

Table 1
Data description.

Sequence name	*MCG*	*GVH*	*LIP*	*CHG*	*AAC*	*ALM*1	*ALM*2	*Class*
x_1(*FTSN*)	0.00	0.51	0.48	0.50	0.35	0.67	0.44	im
x_2(*FTSQ*)	0.10	0.49	0.48	0.50	0.41	0.67	0.21	im
x_3(*MOTB*)	0.30	0.51	0.48	0.50	0.42	0.61	0.34	im
x_4(*TOLA*)	0.61	0.47	0.48	0.50	0.00	0.80	0.32	im
x_5(*TOLQ*)	0.63	0.75	0.48	0.50	0.64	0.73	0.66	im
x_6(*EMRB*)	0.71	0.52	0.48	0.50	0.64	1.00	0.99	im
x_7(*ATKC*)	0.85	0.53	0.48	0.50	0.53	0.52	0.35	imS
x_8(*NFRB*)	0.63	0.49	0.48	0.50	0.54	0.76	0.79	imS
x_9(*NLPA*)	0.75	0.55	1.00	1.00	0.40	0.47	0.30	imL
x_{10}(*CYOA*)	0.70	0.39	1.00	0.50	0.51	0.82	0.84	imL

$$[x_1]_{C^h} = \frac{1}{x_1} + \frac{0.5353}{x_2} + \frac{0.2941}{x_3} + \frac{0}{x_4} + \frac{0}{x_5} + \frac{0}{x_6} + \frac{0}{x_7} + \frac{0}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$[x_2]_{C^h} = \frac{0.5353}{x_1} + \frac{1}{x_2} + \frac{0.5294}{x_3} + \frac{0}{x_4} + \frac{0}{x_5} + \frac{0}{x_6} + \frac{0}{x_7} + \frac{0}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$[x_3]_{C^h} = \frac{0.2941}{x_1} + \frac{0.5294}{x_2} + \frac{1}{x_3} + \frac{0}{x_4} + \frac{0.2235}{x_5} + \frac{0}{x_6} + \frac{0}{x_7} + \frac{0.0909}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$[x_4]_{C^h} = \frac{0}{x_1} + \frac{0}{x_2} + \frac{0}{x_3} + \frac{1}{x_4} + \frac{0}{x_5} + \frac{0}{x_6} + \frac{0}{x_7} + \frac{0}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$[x_5]_{C^h} = \frac{0}{x_1} + \frac{0}{x_2} + \frac{0.2235}{x_3} + \frac{0}{x_4} + \frac{1}{x_5} + \frac{0.3333}{x_6} + \frac{0.3737}{x_7} + \frac{0.3067}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$[x_6]_{C^h} = \frac{0}{x_1} + \frac{0}{x_2} + \frac{0}{x_3} + \frac{0}{x_4} + \frac{0.3334}{x_5} + \frac{1}{x_6} + \frac{0}{x_7} + \frac{0.52}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$[x_7]_{C^h} = \frac{0}{x_1} + \frac{0}{x_2} + \frac{0}{x_3} + \frac{0}{x_4} + \frac{0.3737}{x_5} + \frac{0}{x_6} + \frac{1}{x_7} + \frac{0.1111}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$[x_8]_{C^h} = \frac{0}{x_1} + \frac{0}{x_2} + \frac{0.0909}{x_3} + \frac{0}{x_4} + \frac{0.3067}{x_5} + \frac{0.5200}{x_6} + \frac{0.1111}{x_7} + \frac{1}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$[x_9]_{C^h} = \frac{0}{x_1} + \frac{0}{x_2} + \frac{0}{x_3} + \frac{0}{x_4} + \frac{0}{x_5} + \frac{0}{x_6} + \frac{0}{x_7} + \frac{0}{x_8} + \frac{1}{x_9} + \frac{0}{x_{10}},$$
$$[x_{10}]_{C^h} = \frac{0}{x_1} + \frac{0}{x_2} + \frac{0}{x_3} + \frac{0}{x_4} + \frac{0}{x_5} + \frac{0}{x_6} + \frac{0}{x_7} + \frac{0}{x_8} + \frac{0}{x_9} + \frac{1}{x_{10}},$$
$$Y_1 = \{x_1, x_2, x_3, x_4, x_5, x_6\},$$
$$Y_2 = \{x_7, x_8\},$$
$$Y_3 = \{x_9, x_{10}\},$$

where Y_1, Y_2, Y_3 are the decision classes induced by decision attribute.

Furthermore, we can obtain the upper and lower approximations of Y_1:

$$\underline{C^h}(Y_1) = \frac{1}{x_1} + \frac{1}{x_2} + \frac{0.9091}{x_3} + \frac{1}{x_4} + \frac{0.6263}{x_5} + \frac{0.4800}{x_6} + \frac{0}{x_7} + \frac{0}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$\overline{C^h}(Y_1) = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \frac{1}{x_4} + \frac{1}{x_5} + \frac{1}{x_6} + \frac{0.3737}{x_7} + \frac{0.5200}{x_8} + \frac{0}{x_9} + \frac{0}{x_{10}},$$
$$\underline{HC^h}(Y_1) = \{x_1, x_2, x_4\},$$
$$\overline{HC^h}(Y_1) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}.$$

Thus, it is easy to know that

$$(\underline{C^h}(Y_1))_1 = \{x_1, x_2, x_4\} = \underline{HC^h}(Y_1),$$
$$(\overline{C^h}(Y_1))_0 = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\} = \overline{HC^h}(Y_1).$$

The results are in accordance with Theorem 5.8.

Suppose that parameter $\delta = 0.3$. One can obtain that $\delta_{C^h}(x_1) = \{x_1, x_2\}$, $\delta_{C^h}(x_2) = \{x_1, x_2, x_3\}$, $\delta_{C^h}(x_3) = \{x_2, x_3\}$, $\delta_{C^h}(x_4) = \{x_4\}$, $\delta_{C^h}(x_5) = \{x_5\}$, $\delta_{C^h}(x_6) = \{x_6, x_8\}$, $\delta_{C^h}(x_7) = \{x_7\}$, $\delta_{C^h}(x_8) = \{x_6, x_8\}$, $\delta_{C^h}(x_9) = \{x_9\}$, $\delta_{C^h}(x_{10}) = \{x_{10}\}$. Thus

$$\underline{NC^h}_{0.3}(Y_1) = \{x_1, x_2, x_3, x_4, x_5\},$$
$$\{x_i | x_j \in Y_1 \text{ if } d_{C^n}(x_i, x_j) = 0.3 \wedge d_{C^c}(x_i, x_j) = 0\} = \emptyset, \text{ and}$$
$$\overline{NC^h}_{0.3}(Y_1) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_8\}.$$

Because the relationships between λ and δ in lower and upper approximations are $\delta = f^{-1}(1 - \lambda)$ and $\delta = f^{-1}(\lambda)$, obtaining that $\lambda = 1 - (1 - 2 \times 0.3) = 0.6$ and $\lambda = 1 - 2 \times 0.3 = 0.4$, respectively. Therefore, we can obtain that

$$(\underline{C^h}(Y_1))_{0.6} = \{x_1, x_2, x_3, x_4, x_5\} = \underline{NC^h}_{0.3}(Y_1) - \emptyset,$$
$$(\overline{C^h}(Y_1))_{0.4} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_8\} = \overline{NC^h}_{0.3}(Y_1).$$

The given example easily explains Theorems 5.9 and 5.10.

In addition, suppose that $\beta = 0.6$; it is easy to obtain that

$$\underline{WC^h}_{0.6}(Y_1) = \{x_1, x_2, x_3, x_4, x_5\} = (\underline{C^h}(Y_1))_{0.6},$$
$$\overline{WC^h}_{0.6}(Y_1) = \{x_1, x_2, x_3, x_4, x_5, x_6\} = (\underline{C^h}(Y_1))_{0.6}.$$

The above equations expound and illustrate Theorem 5.11.

6. Conclusions

This paper clarifies the relationships among the generalized rough set models for hybrid data. To approach the target, we investigated the relationships among the rough sets from two viewpoints: constructing information granules and rough approximations. We first investigated in detail the construction of fuzzy and neighborhood hybrid granules. We then analyzed the relationships among these rough approximations. We came to the following conclusions: Hu's fuzzy rough approximations are special cases of both neighborhood and Wang's fuzzy rough approximations. One-to-one correspondence relationships exist between Wang's fuzzy and neighborhood rough approximations. Wang's fuzzy and neighborhood rough approximations are the cut sets of Dubois' and Radzikowska and Kerre's fuzzy rough approximations, respectively. These results can help researchers both understand these generalized rough sets and select a proper model for a given application.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 71031006, 70971080, 60903110), Special prophase project for the National Key Basic Research and Development Program of China (973) (No. 2011CB311805), the Foundation of Doctoral Program Research of the Ministry of Education of China (20101401110002) and the Natural Science Foundation of Shanxi Province (Nos. 2009021017-1, 2010021017-3).

References

[1] J. Bazan, J.F. Peters, A. Skowron, H.S. Nguyen, M. Szczuka, Rough set approach to pattern extraction from classifiers, Electronic Notes in Theoretical Computer Science 82 (4) (2003) 1–10.
[2] R. Bhatt, M. Gopal, On fuzzy-rough sets approach to feature selection, Pattern Recognition Letters 26 (2005) 965–975.
[3] D.G. Chen, Q.H. Hu, Y.P. Yang, Parameterized attribute reduction with Gaussian kernel based fuzzy rough sets, Information Sciences 181 (23) (2011) 5169–5179.
[4] D.G. Chen, S.Y. Zhao, Local reduction of decision system with fuzzy rough sets, Fuzzy Systems and Sets 161 (2010) 1871–1883.
[5] C. Cornelis, R. Jensen, G. Hurtado, D. Ślęzak, Attribute selection with fuzzy decision reducts, Information Sciences 180 (2) (2010) 209–224.
[6] R.K. De, J. Basak, S.K. Pal, Neuro-fuzzy feature evaluation with theoretical analysis, Neural Networks 12 (10) (1999) 1429–1455.
[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] D. Dubois, H. Prade, Twofold fuzzy sets and rough sets-some issues in knowledge representation, Fuzzy Sets and Systems 23 (1987) 3–18.
[9] 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.
[10] 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.
[11] 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.
[12] Q.H. Hu, J.F. Liu, C.X. Wu, Neighborhood rough set based heterogeneous feature subset selection, Information Sciences 178 (18) (2008) 3577–3594.
[13] Q.H. Hu, S. An, D.R. Yu, Soft fuzzy rough sets for robust feature evaluation and selection, Information Sciences 180 (22) (2010) 4384–4400.
[14] J.S.R. Jang, C.T. Sun, E. Mizutani, Neuro-fuzzy and Soft Computing – A Computational Approach to Learning and Machine Intelligence, Prentice Hall, New Jersey, 1997.
[15] R. Jensen, Q. Shen, Fuzzy-rough sets for descriptive dimensionality reduction, in: Proceedings of IEEE International Conference on Fuzzy Systems, 2002, pp. 29–34.
[16] R. Jensen, C. Cornelis, Fuzzy-rough nearest neighbour classification and prediction, Theoretical Computer Science 412 (2011) 5871–5884.

[17] G.J. Klir, B. Yuan, Fuzzy Logic: Theory and Applications, Prentice-Hall, Englewood Cliffs, NJ, 1995.
[18] D.Y. Li, B. Zhang, Y. Leung, On knowledge reduction in inconsistent decision information systems, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 12 (5) (2004) 651–672.
[19] T.Y. Lin, Neighborhood systems-application to qualitative fuzzy and rough sets, in: P.P. Wang (Ed.), Advances in Machine Intelligence and Soft-computing, Department of Electrical Engineering, Duke University Durham, North Carolina, USA, 1997, pp. 132–155.
[20] 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.
[21] J.Y. Liang, Y.H. Qian, Information granules and entropy theory in information systems, Science in China (Series F) 51 (9) (2008) 1427–1444.
[22] J.Y. Liang, J. H Wang, Y.H. Qian, A new measure of uncertainty based on knowledge granulation for rough sets, Information Sciences 179 (2009) 458–470.
[23] G.L. Liu, Y. Sai, Invertible approximation operators of generalized rough sets and fuzzy rough sets, Information Sciences 180 (11) (2010) 2221–2229.
[24] A.M. Martinez, A.C. Kak, PCA versus LDA, IEEE Transaction on Pattern Analysis and Machine Intelligence 23 (2) (2001) 228–233.
[25] J.S. Mi, W.X. Zhang, An axiomatic characterization of a fuzzy generalization of rough sets, Information Sciences 160 (2004) 235–249.
[26] N.N. Morsi, M.M. Yakout, Axiomatics for fuzzy rough sets, Fuzzy Sets and Systems 100 (1–3) (1998) 327–342.
[27] Y. Ouyang, Z.D. Wang, H.P. Zhang, On fuzzy rough sets based on tolerance relations, Information Sciences 180 (4) (2010) 532–542.
[28] Z. Pawlak, Rough Sets, Theoretical Aspects of Reasoning about Data, Kluwer Academic Publisher, London, 1991.
[29] W. Pedrycz, G. Vukovich, Feature analysis through information granulation and fuzzy sets, Pattern Recognition 35 (2002) 825–834.
[30] Y.H. Qian, C.Y. Dang, J.Y. Liang, D.W. Tang, Set-valued ordered information systems, Information Sciences 179 (2009) 2809–2832.
[31] 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.
[32] Y.H. Qian, J.Y. Liang, Y.Y. Yao, C.Y. Dang, MGRS: a multi-granulation rough set, Information Sciences 180 (2010) 949–970.
[33] A.M. Radzikowska, E.E. Kerre, A comparative study on fuzzy-rough sets, Fuzzy Sets and Systems 26 (2002) 137–155.
[34] R. Jensen, Q. Shen, Fuzzy-rough sets assisted attribute selection, IEEE Transactions on Fuzzy systems 15 (1) (2007) 73–89.
[35] Q. Shen, R. Jensen, Selecting informative features with fuzzy-rough sets and its application for complex systems monitoring, Pattern Recognition 37 (7) (2004) 1351–1363.
[36] A. Skowron, C. Rauszer, The Discernibility Matrices and Functions in Information Systems, Intelligent Decision Support: Handbook of Applications and Advances of Rough Set Theory (1992) 331–362.
[37] V. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.
[38] X.Z. Wang, E.C.C. Tsang, S.Y. Zhao, D.G. Chen, D.S. Yeung, Learning fuzzy rules from fuzzy samples based on rough set technique, Information Sciences 177 (2007) 4493–4514.
[39] 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.
[40] W. Wei, J.Y. Liang, Y.H. Qian, F. Wang, An attribute reduction approach and its accelerated version for hybrid data, in: The 8th IEEE International Conference on Cognitive Informatics, 2009, pp. 167–173.
[41] W.Z. Wu, W.X. Zhang, Neighborhood operator systems and approximations, Information Sciences 144 (1–4) (2002) 201–217.
[42] W.Z. Wu, W.X. Zhang, Constructive and axiomatic approaches of fuzzy approximation operators, Information Sciences 159 (3–4) (2004) 233–254.
[43] Y.Y. Yao, Two views of the theory of rough sets in finite universes, International Journal of Approximate Reasoning 15 (1996) 291–317.
[44] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Information Sciences 111 (1–4) (1998) 239–259.
[45] Y.Y. Yao, Neighborhood systems and approximate retrieval, Information Sciences 176 (23) (2006) 3431–3452.
[46] 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.
[47] W. Ziarko, Variable precision rough set model, Journal of Computer and System Science 46 (1993) 39–59.