Information Sciences 180 (2010) 2003–2022

Contents lists available at ScienceDirect

Information Sciences

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

Fuzzy preference based rough sets

Qinghua Hu *, Daren Yu, Maozu Guo

Harbin Institute of Technology, Harbin 150001, PR China

ARTICLE INFO

Article history:
Received 2 July 2008
Received in revised form 7 November 2009
Accepted 12 January 2010

Keywords:
Preference analysis
Multi-criteria decision-making
Fuzzy sets
Rough sets
Dependency

ABSTRACT

Preference analysis is an important task in multi-criteria decision making. The rough set theory has been successfully extended to deal with preference analysis by replacing equivalence relations with dominance relations. The existing studies involving preference relations cannot capture the uncertainty presented in numerical and fuzzy criteria. In this paper, we introduce a method to extract fuzzy preference relations from samples characterized by numerical criteria. Fuzzy preference relations are incorporated into a fuzzy rough set model, which leads to a fuzzy preference based rough set model. The measure of attribute dependency of the Pawlak's rough set model is generalized to compute the relevance between criteria and decisions. The definitions of upward dependency, downward dependency and global dependency are introduced. Algorithms for computing attribute dependency and reducts are proposed and experimentally evaluated by using two publicly available data sets.

1. Introduction

Many decision-making problems are characterized by the ranking of objects according to a set of criteria with pre-defined preference-ordered decision classes, such as credit approval, stock risk estimation, university ranking, teaching evaluation, etc. [2,6,34,40,47,51]. Models and algorithms were proposed for extracting and aggregating fuzzy preference relations based on distinct criteria [41,42]. The underlying objectives are to understand the decision process, to build decision models and to learn decisions rules from data. As available information is usually obtained from different evaluating criteria or experts, the derived decisions may be inconsistent. The combination of multiple sources of information is a challenge for decision analysis.

The rough set theory provides an effective tool for dealing with inconsistency and incomplete information [4,9,10,20,37–39]. It has been widely applied in dependency analysis [3,22–24], feature evaluation [25,26,46], attribute reduction [48,49,54] and rule extraction [3,40,44,45,49]. The Pawlak's rough set model is constructed based on equivalence relations. These relations are viewed by many to be one of the main limitations when employing the model to complex decision tasks [35,36,53]. In multiple criteria decision-making problems, there are preference structures between conditions and decisions. Greco et al. introduced a dominance rough set model that is suitable for preference analysis [11–13,16,17,42]. They examined the decision-making problem with multiple attributes and multiple criteria, where dominance relations were extracted from multiple criteria and similarity relations were constructed from numerical attributes and equivalence relations were constructed from nominal features. An extensive review of multi-criteria decision analysis based on dominance rough sets is given in Greco et al. [14]. Dominance rough sets have also been applied to ordinal attribute reduction [24,28] and multi-criteria classification [2]. Recently, fuzzy set extension of dominance rough set approach [12,14,15,18] and dominance rough sets based on rough-graded preference relations [44] were proposed. In these new models, one not only knows that an action

* Corresponding author.
E-mail address: huqinghua@hit.edu.cn (Q. Hu).

0020-0255/$ - see front matter
doi:10.1016/j.ins.2010.01.015

x is preferred to another action y, but also has information on how much x is preferred to action y. The characterization of graded preference decisions is finer than the original ones. There are differences between graded preference relations and fuzzy preference relations, which has been widely discussed in multi-criteria decision. Fuzzy preference relations can reflect the degree of preference quantitatively, while preference relations can only represent qualitative information. In this paper, we introduce a fuzzy rough set model based on fuzzy preference relations. This model can be considered as a special case of the model presented in Greco et al. [18].

As claimed in many studies, fuzzy sets and rough sets are complementary rather than competitive in dealing with uncertainty [7,52]. Fuzzy sets deal with the uncertainty in representing concepts, while rough sets characterize inconsistency in describing new concepts in terms of definable concepts. As pointed out by Zadeh [56,57], in human reasoning and concept formation, the granules used are fuzzy rather than Boolean. It is therefore useful to integrate rough sets with fuzzy sets into a unified framework. The first fuzzy rough model was constructed by Dubois and Prade [8], where fuzzy similarity relations satisfying reflexivity, symmetry and max–min transitivity were used. The *min* and *max* fuzzy set-theoretic operators were used in the definition of the fuzzy lower approximation and fuzzy upper approximation. The operator *min* and its dual operator *max* are one pair of special t-norm and t-conorm [27]. Radzikowska and Kerre gave a more general definition of fuzzy rough sets based on the T-equivalence relation [41]. They defined a family of fuzzy rough sets based on fuzzy similarity relations. The fuzzy lower approximation and upper approximation are determined by an implicator and a t-norm, respectively. Morsi and Yakout [33] used t-norms and T-residuated implication to define fuzzy rough sets. Mi and Zhang [32] proposed a new definition of fuzzy rough sets based on a residual implication θ and its dual σ, where fuzzy similarity relations were employed to generate fuzzy granulated spaces. Fuzzy rough sets have successfully been applied to many fields, such as feature evaluation [22,25], attribute reduction [19,25,26], rule extraction [21], classification tree induction [1,29], medical analysis [19], stock prediction [46] and case based reasoning [10,31].

The fuzzy rough set models constructed based on fuzzy similarity relations or T-similarity relations are incapable of analyzing data with preference structures. Greco et al. discussed fuzzy dominance relations in rough set theory. Wu et al. proposed a generalized fuzzy rough set model based on general fuzzy relations and max-min fuzzy operators [50]. Mi and Zhang presented a fuzzy rough set model based on fuzzy residual implication θ and its dual σ [32]. In [55], Yeung et al. reviewed the previous work and showed two classes of models capable of defining fuzzy rough sets based on arbitrary fuzzy relations.

Fuzzy preference relations, as a special class of fuzzy relations, can be integrated with these general fuzzy rough set models. A fuzzy preference rough set model can be established for fuzzy preference analysis. Greco et al. introduced fuzzy dominance rough sets using possibility and necessity measures [11]. They did not discuss how to generate fuzzy dominance relations from data. In this paper, we will introduce an algorithm to generate fuzzy preference relations from data and integrate the fuzzy relations with general fuzzy rough sets, and propose a general fuzzy preference based rough set model. Here we will construct algorithms for analyzing the dependency and significance of criteria. The lower and upper approximations in fuzzy preference analysis can be interpreted as the pessimistic and optimistic decisions, respectively.

The paper is organized as follows. We present preliminary knowledge about rough sets, dominance rough sets and fuzzy rough sets in Section 2. We describe algorithms used to extract fuzzy preference relations from samples and discuss their properties in Section 3. Section 4 presents the definitions of fuzzy preference approximations. Section 5 discusses the measures of approximation quality and algorithms for criteria reduction. Experimental analysis is provided in Section 6.

2. Preliminaries

2.1. Rough sets

An information system is a pair $IS = (U, A)$, where $U = \{x_1, \ldots, x_n\}$ is a nonempty finite set of objects and $A = \{a_1, a_2, \ldots, a_L\}$ is a finite set of attributes used to characterize the objects. For a subset of attributes $B \subseteq A$, we define a binary relation $IND(B)$, called the$B$-indiscernibility relation, as follows: $IND(B) = \{(x, y) \in U \times U : f(x, a) = f(y, a), \forall a \in B\}$, where $f(x, a)$ is the feature value of object x. The relation $IND(B)$ is an equivalence relation with $IND(B) = \bigcap_{a \in B} IND(\{a\})$. The indiscernibility relation divides the objects into a family of disjoint subsets. The set $[x]_B$ denotes the equivalence class containing x. For an arbitrary subset$X \subseteq U$, two subsets of objects, computed as $\underline{B}X = \bigcup\{[x]_B : [x]_B \subseteq X\}$ and $\overline{B}X = \bigcup\{[x]_B : [x]_B \bigcap X \neq \emptyset\}$, are called B-lower and B-upper approximations, respectively. X is said to be definable if $\underline{B}X = \overline{B}X$; otherwise, X is a rough set. We call $BN(X) = \overline{B}X - \underline{B}X$ the boundary of X.

Consider a decision table $DT = (U, A = C \bigcup D)$, where the set of attributes are divided into condition C and decision D. In classification analysis, we have the problem of using the concepts generated by condition attributes to approximately describe the decision. Assume that the objects are grouped into N decision classes, denoted by $d_1, d_2, \ldots, d_N$. Then $\underline{B}d_i = \bigcup\{[x]_B : [x]_B \subseteq d_i\}$ and $\overline{B}d_i = \bigcup\{[x]_B : [x]_B \bigcap d_i \neq \emptyset\}$. Correspondingly, B-lower and B-upper approximations of decision D are defined as

$$\underline{B}D = \bigcup_{i=1}^{N} \underline{B}d_i \quad \text{and} \quad \overline{B}D = \bigcup_{i=1}^{N} \overline{B}d_i,$$

respectively.

The approximation quality of D with respect to B is computed by

$$\gamma_B(D) = \frac{|\bigcup_{i=1}^{N} \underline{B}d_i|}{|U|}.$$

The approximation quality is also called dependency, reflecting relevance between condition and decision.

2.2. Dominance rough sets

In multi-criteria decision analysis, there is a preferential structure on the decision classes. Without loss of generality, we assume that $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_N$. Given a set of samples, one requires knowledge about the decisions $x \geqslant_d y$ if $x \geqslant_B y$ or $x \leqslant_d y$ if $x \leqslant_B y$, where x and y are two objects, $\geqslant_d, \geqslant_B, \leqslant_d$ and $\leqslant_B$ stand for relations of less than and not greater than in terms of d and B, respectively. Preference decision and classification are two different types of problems. In classification, the objects with the same condition values should be classified into the same decision class; otherwise, the classification is inconsistent. In preference analysis, x's decision should not be worse than y's if x is better than y in terms of criterion a; otherwise, the decision is inconsistent.

For preference analysis, equivalence relations in Pawlak's rough sets are replaced with dominance relations. Assumed that the criteria consist of ordinal discrete or numerical values, the following sets are defined. For $x \in U, X \subseteq U, a \in A, B \subseteq A$,

$$[x]_a^{\geqslant} = \{y : y \geqslant_a x\}, \quad [x]_a^{\leqslant} = \{y : y \leqslant_a x\};$$
$$[x]_B^{\geqslant} = \{y : y \geqslant_a x, \forall a \in B\}, \quad [x]_B^{\leqslant} = \{y : y \leqslant_a x, \forall a \in B\};$$
$$[X]_B^{\leqslant} = \bigcup_{x \in X} [x]_B^{\leqslant}, [X]_B^{\geqslant} = \bigcup_{x \in X} [x]_B^{\geqslant};$$

The upward and downward approximations of $X \subseteq U$ in terms of criteria B are defined as

(1) upward approximation: $\underline{B^{\geqslant}}X = \{x : [x]_B^{\geqslant} \subseteq X\}$ and $\overline{B^{\geqslant}}X = \{x : [x]_B^{\leqslant} \bigcap X \neq \emptyset\}$;
(2) downward approximation: $\underline{B^{\leqslant}}X = \{x : [x]_B^{\leqslant} \subseteq X\}$ and $\overline{B^{\leqslant}}X = \{x : [x]_B^{\geqslant} \bigcap X \neq \emptyset\}$.

As to decision d_i, we construct two subsets of objects, denoted by $d_i^{\geqslant}$ and $d_i^{\leqslant}$ as $d_i^{\geqslant} = \bigcup_{j \geqslant i} d_j$ and $d_i^{\leqslant} = \bigcup_{j \leqslant i} d_j$. The lower and upper approximations of these sets are given by

(1) upward approximation: $\underline{B^{\geqslant}}d_i^{\geqslant} = \{x : [x]_B^{\geqslant} \subseteq d_i^{\geqslant}\}$ and $\overline{B^{\geqslant}}d_i^{\geqslant} = \{x : [x]_B^{\leqslant} \bigcap d_i^{\geqslant} \neq \emptyset\}$;
(2) downward approximation: $\underline{B^{\leqslant}}d_i^{\leqslant} = \{x : [x]_B^{\leqslant} \subseteq d_i^{\leqslant}\}$ and $\overline{B^{\leqslant}}d_i^{\leqslant} = \{x : [x]_B^{\geqslant} \bigcap d_i^{\leqslant} \neq \emptyset\}$.

2.3. Fuzzy rough sets

The concept of fuzzy rough sets was proposed and studied by many authors [5,8,23,30]. A detailed survey on this topic can be found in Yeung et al. [55]. We briefly review some of the basic concepts and notions.

Let U be a nonempty universe, R an arbitrary fuzzy relation on U. For a fuzzy subset A, the lower and upper approximations of A are defined as

(1) S-lower approximation operator: $\underline{R_S}A(x) = \inf_{u \in U} S(N(R(x,u)), A(u))$;
(2) T-upper approximation operator: $\overline{R_T}A(x) = \sup_{u \in U} T(R(x,u), A(u))$;
(3) ϑ-lower approximation operator: $\underline{R_\vartheta}A(x) = \inf_{u \in U} \vartheta(R(x,u), A(u))$;
(4) σ-upper approximation operator: $\overline{R_\sigma}A(x) = \sup_{u \in U} \sigma(N(R(x,u)), A(u))$;

where S, T, ϑ and σ are defined as follows.

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

(1) *min* operator: $T_M(x,y) = \min\{x,y\}$;
(2) algebraic product: $T_P(x,y) = x \cdot y$;
(3) bold intersection:$T_L(x,y) = \max\{0, x+y-1\}$.

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

(1) *max* operator $S_M(x,y) = \max\{x,y\}$;
(2) probabilistic sum $S_P(x,y) = x + y - x \cdot y$;
(3) bounded sum $S_L(x,y) = \min\{1, x+y\}$.

A negator N is a decreasing mapping $[0,1] \to [0,1]$ satisfying $N(0) = 1$ and $N(1) = 0$. The negator $N_s(x) = 1 - x$ is referred to as the standard negator. A negator N is involutive if $N(N(x)) = x$ for all$x \in [0,1]$, an involutive negator is continuous and strictly decreasing. Given a negator N, a t-norm T and a t-conorm S are dual with respect to N if the De Morgan laws are satisfied, i.e., $S(N(x), N(y)) = N(T(x,y)), T(N(x), N(y)) = N(S(x,y))$.

For $A \in F(U)$, where $F(U)$ is the fuzzy power set of U, the symbol co_N denotes the fuzzy complement of A determined by a negator N, i.e. for every $x \in U, (co_N A)(x) = N(A(x))$. Given a triangular norm T, the binary operation on I, $\vartheta_T(\alpha, \gamma) = \sup\{\theta \in I : T(\alpha, \theta) \leqslant \gamma\}$, $\alpha, \gamma \in I$, is called a implicator related to . If T is lower semi-continuous, ϑ_T is called the T-residual implication. The properties of T-residual implication ϑ_T are listed in Yeung et al. [55].

In [32], σ is defined as $\sigma(a,b) = \inf\{c \in [0,1] : S(a,c) \geqslant b\}$ for a t-conorm S. If T and S are dual with respect to an involutive negator N, ϑ and σ are dual with respect to the involutive negator N, i.e., $\sigma(N(a), N(b)) = N(\vartheta(a,b)), \vartheta(N(a), N(b)) = \sigma(N(a,b))$.

3. Fuzzy preference relations and fuzzy preference granules

In real world applications of fuzzy rough set models, three key issues must be addressed: inducing a granular structure on the universe based on an attribute or a criterion, aggregating the granular structures deduced from different attributes or criteria, and computing the lower and upper approximations of decisions. In this section, we will introduce a technique to compute the fuzzy preference relations.

There are two kinds of preference relations widely used in many decision-making models [20].

(1) multiplicative preference relations: a multiplicative preference $R \in U \times U$ is represented by a relation matrix $(r_{ij})_{n \times n}$, where r_{ij} is interpreted as the ratio of the preference degree of x_i to that of x_j, i.e., x_i is r_{ij} times as good as x_j. It was suggested that $r_{ij} \in \{1, 2, \ldots, n\}$. In this case, the fuzzy preference relation R is usually assumed to be a multiplicative reciprocal, i.e., $r_{ij} \cdot r_{ji} = 1, \forall i,j \in \{1, 2, \ldots, n\}$.

(2) Fuzzy preference relations: a fuzzy preference relation R is a fuzzy set on the product set $U \times U$, which is characterized by a membership function $\mu_R : U \times U \to [0,1]$. If the cardinality of U is finite, the fuzzy preference relation can also be represented by a $n \times n$ matrix $(r_{ij})_{n \times n}$, where r_{ij} is interpreted as the preference degree of x_i over $x_j : r_{ij} = 1/2$ indicates there is no difference between x_i and $x_j; r_{ij} > 1/2$ shows x_i is preferred to x_j and $r_{ij} = 1$ means x_i is absolutely preferred to x_j. On the other hand, $r_{ij} < 1/2$ indicates x_j is preferred to x_i. In this case, the preference matrix is usually assumed to be an additive reciprocal, i.e., $r_{ij} + r_{ji} = 1, \forall i,j \in \{1, 2, \ldots, n\}$.

In applications, preference structures are usually represented with criteria characterized by a set of ordinal discrete values or numerical values.

Example 1. Table 1 shows 10 manuscripts described by two criteria: originality and writing quality, denoted by a_1 and a_2, respectively. D is the decision of the manuscripts. The decision values of these manuscripts are *accept*, *revise* and *reject*, denoted by 3, 2 and 1. The task is to analyze the consistency of the decision and to compute the dependency between each criterion and decision.

Given a universe of finite objects $U = \{x_1, \ldots, x_n\}$, a is a numerical feature which describes the objects. The feature value of x is $f(x,a)$. The upward and downward fuzzy preference relations over U are computed by

$$r_{ij}^{>} = \frac{1}{1 + e^{-k(f(x_i,a) - f(x_j,a))}} \quad \text{and} \quad r_{ij}^{<} = \frac{1}{1 + e^{k(f(x_i,a) - f(x_j,a))}},$$

where k is a positive constant.

$f(x) = \frac{1}{1+e^{-kx}}$ is the well-known Logsig sigmoid transfer function used in neural networks. The curves of the function are shown in Fig. 1 with x ranging from -1 to 1, where k = -5, -10, -50, 5, 10, 50, respectively.

If k = -5, -10, -50, we can interpret the curves as membership functions of objects much less than zero; if k = 5, 10 or 50, the curves are the membership functions of objects much greater than zero. The parameter k is used to control the preference degree by users. If k = 5, the membership function describes the degrees in which objects are strongly better than zero; while k = 50 the membership function gives the degrees in which objects are slightly better than zero. In applications, k should be specified by users according to their preferences.

The fuzzy preference relations induced by the originality a_1 and the writing quality a_2 in Example 1 are presented in Tables 2 and 3, respectively, where k = 10.

Table 1
Samples of multi-criteria decision-making.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
a_1	0.28	0.25	0.60	0.48	0.42	0.55	0.78	0.75	0.83	0.85
a_2	0.28	0.31	0.42	0.47	0.51	0.58	0.71	0.78	0.80	0.91
D	1	1	1	2	2	2	2	3	3	3

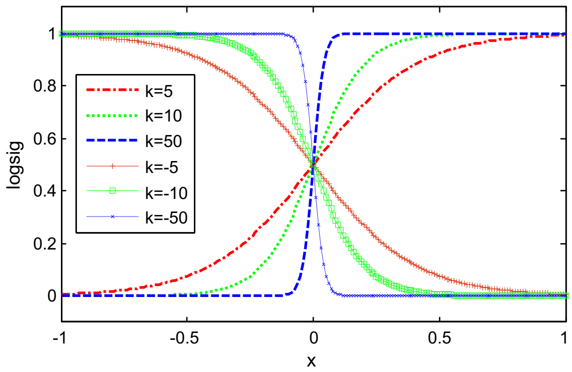

Fig. 1. Curve of function logsig in interval $[-1,1]$.

Table 2
Fuzzy preference relation computed with originality a_1.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	0.50	0.43	0.96	0.88	0.8	0.94	0.99	0.99	1.00	1.00
x_2	0.57	0.50	0.97	0.91	0.85	0.95	1.00	0.99	1.00	1.00
x_3	0.04	0.03	0.50	0.23	0.14	0.38	0.86	0.82	0.91	0.92
x_4	0.12	0.09	0.77	0.50	0.35	0.67	0.95	0.94	0.97	0.98
x_5	0.20	0.15	0.86	0.65	0.50	0.79	0.97	0.96	0.98	0.99
x_6	0.06	0.05	0.62	0.33	0.21	0.50	0.91	0.88	0.94	0.95
x_7	0.01	0.00	0.14	0.05	0.03	0.09	0.50	0.43	0.62	0.67
x_8	0.01	0.01	0.18	0.06	0.04	0.12	0.57	0.50	0.69	0.73
x_9	0.00	0.00	0.09	0.03	0.02	0.06	0.38	0.31	0.50	0.55
x_{10}	0.00	0.00	0.08	0.02	0.01	0.05	0.33	0.27	0.45	0.50

Table 3
Fuzzy preference relation computed with writing a_2.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	0.50	0.57	0.80	0.87	0.91	0.95	0.99	0.99	0.99	1.00
x_2	0.43	0.50	0.75	0.83	0.88	0.94	0.98	0.99	0.99	1.00
x_3	0.20	0.25	0.50	0.62	0.71	0.83	0.95	0.97	0.98	0.99
x_4	0.13	0.17	0.38	0.50	0.6	0.75	0.92	0.96	0.96	0.99
x_5	0.09	0.12	0.29	0.40	0.50	0.67	0.88	0.94	0.95	0.98
x_6	0.05	0.06	0.17	0.25	0.33	0.50	0.79	0.88	0.90	0.96
x_7	0.01	0.02	0.05	0.08	0.12	0.21	0.50	0.67	0.71	0.88
x_8	0.01	0.01	0.03	0.04	0.06	0.12	0.33	0.50	0.55	0.79
x_9	0.01	0.01	0.02	0.04	0.05	0.1	0.29	0.45	0.50	0.75
x_{10}	0.00	0.00	0.01	0.01	0.02	0.04	0.12	0.21	0.25	0.50

Fuzzy preference relations not only reflect the fact that an object x_i is greater (less) than object x_j, but also measure how much x_i is greater (less) than x_j. This shows fuzzy preference relations are more powerful in extracting information from fuzzy data than dominance relations.

Note that $r^{\geqslant}_{ii} = r^{\leqslant}_{ii} = 0.5, \frac{1}{1+e^{-x}} = 1$ if $x = +\infty$, and $\frac{1}{1+e^{-x}} = 0$ if $x = -\infty$. These results are consistent with our intuition: $r_{ij} = 0.5$ indicates that there is no difference between x_i and $r^{\geqslant}_{ij} = 1$ indicates that x_i is much greater than x_j. It is easy to note that $\forall i,j \in \{1,2,\ldots,n\}, r^{\geqslant}_{ij} = r^{\leqslant}_{ji}; r^{\geqslant}_{ij} + r^{\geqslant}_{ji} = 1$ and $r^{\leqslant}_{ij} + r^{\leqslant}_{ji} = 1$ if fuzzy preference relations are computed with the logsig function. Fuzzy preference relations are neither reflexive nor symmetric, but they are transitive, i.e., $r_{ii} \neq 1, r_{ij} \neq r_{ji}$ and $\min_y(R(x,y), R(y,z)) \leqslant R(x,z)$.

Generally speaking, there is more than a single criterion used in decision analysis. One can generate a fuzzy preference relation with respect to each criterion. We next require a technique to aggregate preference relations induced by multiple criteria. There are some methods developed to deal with this problem, such as t-norm and weighted aggregation operators [27,43,51].

Definition 1. Given a task of multi-criteria decision-making, x_i and x_j are two arbitrary objects in the universe. If r_{ij} and s_{ij} are the fuzzy preference degrees of x_i over x_j derived from two criteria a and b, respectively, the aggregated preference of a and b is defined as $\min(r_{ij}, s_{ij})$.

Certainly, we can also use other t-norms, such as algebraic product and bold intersection operators.

Definition 2. Fuzzy preference relations produce nested granularity structures on the universe, the fuzzy subsets of objects greater or less than x_i are defined as

$$[x_i]_a^> = \frac{r_{1i}}{x_1} + \frac{r_{2i}}{x_2} + \cdots + \frac{r_{ni}}{x_n} \quad \text{and} \quad [x_i]_a^< = \frac{r_{i1}}{x_1} + \frac{r_{i2}}{x_2} + \cdots + \frac{r_{in}}{x_n},$$

where $r_{ij} = \frac{1}{1+e^{-k(f(x_i,a)-f(x_j,a))}}, k > 0$.

Example 2 (*Continued*). For object x_1 and criterion of the originality in Example 1, the fuzzy preference granule of objects that are better than x_1 is expressed as

$$[x_1]_{a_1}^> = \frac{0.5}{x_1} + \frac{0.43}{x_2} + \frac{0.96}{x_3} + \frac{0.88}{x_4} + \frac{0.80}{x_5} + \frac{0.94}{x_6} + \frac{0.99}{x_7} + \frac{0.99}{x_8} + \frac{1}{x_9} + \frac{1}{x_{10}}.$$

$[x_1]_{a_1}^>$ is a fuzzy set of objects that are far better than x_1 in terms of the originality a_1. As we know, $f(x_1, a_1) = 0.28$ and $f(x_{10}, a_1) = 0.85, f(x_{10}, a_1)$ is far better than $f(x_1, a_1)$ and $\mu_{[x_1]_{a_1}^>}(x_{10}) = 1$. $\mu_{[x_1]_{a_1}^>}(x_{10})$ is the membership of object x_{10} belonging to the fuzzy set $[x_1]_{a_1}^>$. Similarly,

$$[x_1]_{a_2}^> = \frac{0.5}{x_1} + \frac{0.57}{x_2} + \frac{0.80}{x_3} + \frac{0.87}{x_4} + \frac{0.91}{x_5} + \frac{0.95}{x_6} + \frac{0.99}{x_7} + \frac{0.99}{x_8} + \frac{0.99}{x_9} + \frac{1}{x_{10}}.$$

Let $B = \{a_1, a_2\}$, then $[x_1]_B^> = \frac{0.5}{x_1} + \frac{0.43}{x_2} + \frac{0.80}{x_3} + \frac{0.87}{x_4} + \frac{0.80}{x_5} + \frac{0.94}{x_6} + \frac{0.99}{x_7} + \frac{0.99}{x_8} + \frac{0.99}{x_9} + \frac{1}{x_{10}}$. $[x_1]_B^>$ is a fuzzy set of objects that are much better than x_1 in terms of the writing quality and the originality.

We come to the following conclusions:

(1) $x \geqslant_a y \iff [x]_a^> \subseteq [y]_a^> \iff [y]_a^< \subseteq [x]_a^<$;
(2) $x \geqslant_B y \iff [x]_B^> \subseteq [y]_B^> \iff [y]_B^< \subseteq [x]_B^<$.

4. Approximating preference decision with fuzzy preference granules

We now approximate preference decisions with fuzzy granules. We will introduce a general fuzzy rough set model to compute fuzzy preference approximations.

There are the two ways to define fuzzy lower and upper approximations in the literature. One is based on fuzzy relations [12,14,18,32,50,55] and the other is based on fuzzy coverings [5,30]. We use fuzzy relations based fuzzy rough set models. Consider the following fuzzy approximation operators:

- S-lower approximation operator: $\underline{R_S}A(x) = \inf_{u\in U} S(N(R(x,u)), A(u))$;
- T-upper approximation operator: $\overline{R_T}A(x) = \sup_{u\in U} T(R(x,u), A(u))$,

where $N(a) = 1 - a, S = \max$ and $T = \min$. Fuzzy approximation operators can be written as

$$\underline{R}A(x) = \inf_{u\in U} \max(1 - R(x,u), A(u)),$$
$$\overline{R}A(x) = \sup_{u\in U} \min(R(x,u), A(u)).$$

It is easy to obtain similar results for other fuzzy approximation operators.

In multi-criteria decision analysis, the number of decision classes is usually finite. In Example 1, the decision class {accept, revise, reject}. The preference decision forms a nested structure over the universe. Given a preference decision table $\langle U, C \bigcup D\rangle$, the value domain of D is $\{d_1, d_2, \ldots, d_N\}$ and $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_N$. The nested preference decision structure is given by

$$d_i^{\geqslant} = \bigcup_{j\geq i}^{N} d_j \quad \text{or} \quad d_i^{\leqslant} = \bigcup_{j=1}^{i} d_j.$$

Here $d_i^{\geqslant}$ and $d_i^{\leqslant}$ are subsets of U. We have $d_i^{\geqslant} \supseteq d_j^{\geqslant}$ and $d_i^{\leqslant} \subseteq d_j^{\leqslant}$ if $i \leqslant j$ (d_i denotes the set of samples with decision d_i).

Definition 3. Given fuzzy preference relations $R^>$ and $R^<$, preference decision classes $d_i^{\geqslant}$ and $d_i^{\leqslant}$, the memberships of object x to the lower and upper approximations of $d_i^{\geqslant}$ and $d_i^{\leqslant}$ are defined as

- upward fuzzy lower approximation: $LA_2(j) = \underline{S^>} d_j^{\geqslant}(x_i)$,
- upward fuzzy upper approximation: $\overline{R^>} d_i^{\geqslant}(x) = \sup_{u\in U} \min(R^>(x,u), d_i^{\geqslant}(u))$,
- downward fuzzy lower approximation: $\underline{R^<} d_i^{\leqslant}(x) = \inf_{u\in U} \max(1 - R^<(u,x), d_i^{\leqslant}(u))$,
- downward fuzzy upper approximation: $\overline{R^<} d_i^{\leqslant}(x) = \sup_{u\in U} \min(R^<(x,u), d_i^{\leqslant}(u))$,

where $X(x)$ is the membership of object x to the fuzzy subset X. We require that $\underline{R^{>}}d_1^{\geqslant}(x) = 1, \overline{R^{>}}d_1^{\geqslant}(x) = 1$ and $\underline{R^{<}}d_N^{\leqslant}(x) = 1, \overline{R^{<}}d_N^{\leqslant}(x) = 1$, where d_1 and d_N are the worst and the best decision classes, respectively.

There is small difference between Definition 3 and those proposed in Yeung et al. [55]. We replace $R^{>}(x,u)$ and $R^{<}(x,u)$ in [55] with $R^{>}(u,x)$ and $R^{<}(u,x)$ in the definitions of $\underline{R^{>}}d_i^{\geqslant}(x)$ and $\underline{R^{<}}d_i^{\leqslant}(x)$. The two definitions are identical if R is symmetric. Since fuzzy preference relations do not satisfy the property of symmetry, the two definitions are different.

We show how to compute the lower and upper approximations of preference decision classes.

- For $\underline{R^{>}}d_i^{\geqslant}(x)$, by assuming that $u \in d_i^{\geqslant}$, we have $d_i^{\geqslant}(u) = 1, \max(1 - R^{>}(u,x), d_i^{\geqslant}(u)) = 1$. On the other hand, if $u \notin d_i^{\geqslant}$, i.e., $u \in d_{i-1}^{\leqslant}, d_i^{\geqslant}(u) = 0$ and $\max(1 - R^{>}(u,x), d_i^{\geqslant}(u)) = 1 - R^{>}(u,x)$. As $1 - R^{>}(u,x) \leqslant 1$, we have $\underline{R^{>}}d_i^{\geqslant}(x) = \inf_{u \notin d_i^{\geqslant}} 1 - R^{>}(u,x)$. As $1 - R^{>}(u,x) = R^{<}(u,x)$, thus $\underline{R^{>}}d_i^{\geqslant}(x) = \inf_{u \notin d_i^{\geqslant}} R^{<}(u,x)$. If $i = 1, d_1^{<} = \emptyset$. In this case, we have $\underline{R^{>}}d_1^{\geqslant}(x) = 1$.
- Assume that $u \notin d_i^{\geqslant}$, then $d_i^{\geqslant}(u) = 0$ and $\min(R^{>}(x,u), d_i^{\geqslant}(u)) = 0$; otherwise, $u \in d_i^{\geqslant}$, then $d_i^{\geqslant}(u) = 1, \min(R^{>}(x,u), d_i^{\geqslant}(u)) = R^{>}(x,u)$. So $\overline{R^{>}}d_i^{\geqslant}(x) = \sup_{u \in d_i^{\geqslant}} R^{>}(x,u)$.

We can obtain $\underline{R^{<}}d_i^{\leqslant}(x) = \inf_{u \notin d_i^{\leqslant}} R^{>}(u,x)$ and $\overline{R^{<}}d_i^{\leqslant}(x) = \sup_{u \in d_i^{\leqslant}} R^{<}(x,u)$.

Formula $\underline{R^{>}}d_i^{\geqslant}(x) = \inf_{u \notin d_i^{\geqslant}} 1 - R^{>}(u,x)$ indicates that the membership of x to the lower approximation of $d_i^{\geqslant}$ depends on the samples that do not belong to $d_i^{\geqslant}$ and produce the greatest preference degree over x, while $\overline{R^{>}}d_i^{\geqslant}(x) = \sup_{u \in d_i^{\geqslant}} R^{>}(x,u)$ means the membership of x to the upper approximation of $d_i^{\geqslant}$ depends on the samples that belong to $d_i^{\geqslant}$ and have the least preference degree over x. If there is only one criterion to derive the fuzzy preference relation, we can say that $\underline{R^{>}}d_i^{\geqslant}(x)$ is determined by the greatest sample in a preference decision class $d_{i-1}^{\leqslant}$, while $\overline{R^{>}}d_i^{\geqslant}(x)$ depends on the smallest sample within the preference decision class $d_i^{\geqslant}$. A similar interpretation can also be obtained for $\underline{R^{<}}d_i^{\leqslant}(x) = \inf_{u \in d_i^{\leqslant}} 1 - R^{<}(u,x)$ and $\overline{R^{<}}d_i^{\leqslant}(x) = \sup_{u \in d_i^{\leqslant}} R^{<}(x,u)$.

This interpretation is reasonable and consistent with our intuition. If there is a sample u that is greater than x, but the decision of u is worse than x, then x is not consistent according to the assumption of monotonicity that the decision of u should not be worse than that of x if u is greater than x. In this context, the membership of x to the fuzzy lower approximation of $d_i^{\geqslant}$ should be small; otherwise, if all the samples with decisions worse than d_i have criterion values less than that of x, then x is consistent. Correspondingly, $\underline{R^{>}}d_i^{\geqslant}(x)$ is large. If we use the definition in Wu et al. [49], we can obtain the conclusion that $\underline{R^{>}}d_i^{\geqslant}(x) = \inf_{u \notin d_i^{\geqslant}} 1 - R^{>}(x,u) = \inf_{u \notin d_i^{\geqslant}} R^{<}(x,u)$. The conclusion means that we should find the worst sample not belonging to $d_i^{\geqslant}$ for computing $\underline{R^{>}}d_i^{\geqslant}(x)$. It is contradictory to our intuition.

The relation $\underline{R^{>}}d_i^{\geqslant}(x) \leqslant d_i^{\geqslant}(x) \leqslant \overline{R^{>}}d_i^{\geqslant}(x)$ does not hold in the case that all the decisions are consistent. This problem lies in the irreflexivity of fuzzy preference relations. The disadvantage can be overcome with the following model of fuzzy covering based rough sets [30]:

$$\underline{\mathbb{C}_{FR}}(D)(x) = \sup_{C \in \mathbb{C}} T(C(x), \min_{y \in U} I(C(y), D(y))),$$
$$\overline{\mathbb{C}_{FR}}(D)(x) = \inf_{C \in \mathbb{C}} I(C(x), \max_{y \in U} T(C(y), D(y))),$$

where $\mathbb{C}$ is a fuzzy covering induced by the fuzzy preference relation and C is a fuzzy information granule (refer to [30] and Theorem 4.1).

Example 3 (*Continued*). By mapping the objects in Table 1 into an axis, we get several figures. Fig. 2(1) shows the manuscript distribution in terms of the attribute of originality, and Fig. 2(2) is the distribution with respect to the attribute of writing quality. We compute the lower and upper approximations of manuscripts with respect to originality and writing, respectively.

Consider x_6 as an example. By R_a we denote the fuzzy preference relation induced by attribute a. With respect to the originality a_1, we have

$$\underline{R_{a_1}^{>}}d_1^{\geqslant}(x_6) = 1,$$
$$\underline{R_{a_1}^{>}}d_2^{\geqslant}(x_6) = \inf_{u \notin d_2^{\geqslant}} 1 - R_{a_1}^{>}(x_6,u) = 1 - R_{a_1}^{>}(x_6,x_3) = R_{a_1}^{<}(x_6,x_3) = 1 - 0.62 = 0.38,$$

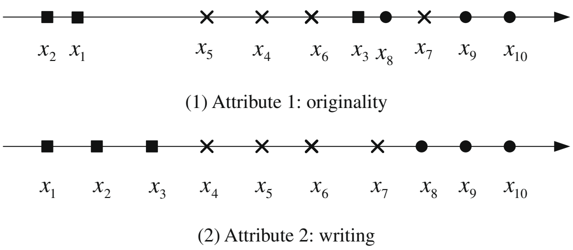

Fig. 2. Consistent and inconsistent decision-making problems.

$$\underline{R^{>}_{a_1}}d_3^{\geqslant}(x_6) = \inf_{u \notin d_3^{\geqslant}} 1 - R^{>}_{a_1}(x_6,u) = 1 - R^{>}_{a_1}(x_6,x_7) = R^{<}_{a_1}(x_6,x_7) = 1 - 0.91 = 0.09,$$

$$\overline{R^{>}_{a_1}}d_1^{\geqslant}(x_6) = \sup_{u \in d_1^{\geqslant}} R^{>}_{a_1}(x_6,u) = R^{>}_{a_1}(x_6,x_2) = 0.95,$$

$$\overline{R^{>}_{a_1}}d_2^{\geqslant}(x_6) = \sup_{u \in d_2^{\geqslant}} R^{>}_{a_1}(x_6,u) = R^{>}_{a_1}(x_6,x_5) = 0.79,$$

$$\overline{R^{>}_{a_1}}d_3^{\geqslant}(x_6) = \sup_{u \in d_3^{\geqslant}} R^{>}_{a_1}(x_6,u) = R^{>}_{a_1}(x_6,x_8) = 0.12.$$

For the writing quality a_2, we have

$$\underline{R^{>}_{a_2}}d_1^{\geqslant}(x_6) = 1,$$

$$\underline{R^{>}_{a_2}}d_2^{\geqslant}(x_6) = \inf_{u \notin d_2^{\geqslant}} 1 - R^{>}_{a_2}(x_6,u) = 1 - R^{>}_{a_2}(x_6,x_3) = R^{<}_{a_2}(x_6,x_3) = 1 - 0.17 = 0.83,$$

$$\underline{R^{>}_{a_2}}d_3^{\geqslant}(x_6) = \inf_{u \notin d_3^{\geqslant}} 1 - R^{>}_{a_2}(x_6,u) = 1 - R^{>}_{a_2}(x_6,x_7) = R^{<}_{a_2}(x_6,x_7) = 1 - 0.79 = 0.21,$$

$$\overline{R^{>}_{a_2}}d_1^{\geqslant}(x_6) = \sup_{u \in d_1^{\geqslant}} R^{>}_{a_2}(x_6,u) = R^{>}_{a_2}(x_6,x_1) = 0.95,$$

$$\overline{R^{>}_{a_2}}d_2^{\geqslant}(x_6) = \sup_{u \in d_2^{\geqslant}} R^{>}_{a_2}(x_6,u) = R^{>}_{a_2}(x_6,x_4) = 0.75,$$

$$\overline{R^{>}_{a_2}}d_3^{\geqslant}(x_6) = \sup_{u \in d_3^{\geqslant}} R^{>}_{a_2}(x_6,u) = R^{>}_{a_2}(x_6,x_8) = 0.12.$$

The lower and upper approximations can be understood as the pessimistic and optimistic decisions in human reasoning, respectively. A pessimistic decision maker thinks that his paper would be accepted with a possibility of 0.5 if it is better than the best rejected paper; while an optimistic decision maker thinks that his paper would be accepted with a possibility of 0.5 if it is better than the worst accepted paper. It is interesting that fuzzy preference based rough sets are consistent with human reasoning.

Given a new submission x11 whose originality is 0.76, its lower and upper assumptions can be calculated based on the expressions given above.

$$\underline{R^{>}_{a_1}}d_2^{\geqslant}(x_{11}) = \inf_{u \notin d_3^{\geqslant}} 1 - R^{>}_{a_1}(x_{11},u) = 1 - R^{>}_{a_1}(x_{11},x_3) = R^{<}_{a_1}(x_{11},x_3) = 1/(1+\exp(10\times(0.60-0.76))) = 0.83,$$

$$\underline{R^{>}_{a_1}}d_3^{\geqslant}(x_{11}) = \inf_{u \notin d_3^{\geqslant}} 1 - R^{>}_{a_1}(x_{11},u) = 1 - R^{>}_{a_1}(x_{11},x_7) = R^{<}_{a_1}(x_{11},x_7) = 1/(1+\exp(10\times(0.78-0.76))) = 0.45,$$

The obtained values show that submission x_{11} possibly requires revision according to the historical samples. However, if another new submission, x12, the originality is 0.9, we will have

$$\underline{R^{>}_{a_1}}d_2^{\geqslant}(x_{12}) = \inf_{u \notin d_3^{\geqslant}} 1 - R^{>}_{a_1}(x_{12},u) = 1 - R^{>}_{a_1}(x_{12},x_3) = R^{<}_{a_1}(x_{12},x_3) = 1/(1+\exp(10\times(0.60-0.90))) = 0.95,$$

$$\underline{R^{>}_{a_1}}d_3^{\geqslant}(x_{12}) = \inf_{u \notin d_3^{\geqslant}} 1 - R^{>}_{a_1}(x_{12},u) = 1 - R^{>}_{a_1}(x_{12},x_7) = R^{<}_{a_1}(x_{12},x_7) = 1/(1+\exp(10\times(0.78-0.90))) = 0.77.$$

In this case, submission x_{12} should be accepted with a preference degree of 0.77 and should receive a decision no less than 'revise' with a preference degree of 0.95.

Given a fuzzy preference decision table $\langle U,C,D\rangle, B \subseteq C, R^{>}$ and $R^{<}$ are fuzzy preference relations generated by criteria B and functions $\frac{1}{1+e^{-k(x_i-x_j)}}$ and $\frac{1}{1+e^{k(x_i-x_j)}}$. $R^{\geqslant}$ and $R^{\leqslant}$ are crisp preference relations generated by criteria B. If $k=+\infty$, and $x_i > x_j, \frac{1}{1+e^{-k(x_i-x_j)}} = 1$ and; otherwise, $\frac{1}{1+e^{k(x_i-x_j)}} = 0$, $x_i < x_j, \frac{1}{1+e^{-k(x_i-x_j)}} = 0$ and $\frac{1}{1+e^{k(x_i-x_j)}} = 1$. In this case, we have $\forall d_i \in D, \underline{R^{>}}d_i^{\geqslant} = \underline{R^{\geqslant}}d_i^{\geqslant}; \underline{R^{<}}d_i^{\leqslant} = \underline{R^{\leqslant}}d_i^{\leqslant}$.

The above analysis shows that fuzzy preference based rough sets will reduce to the crisp preference based rough sets if the gradient coefficient $k=+\infty$. In fact, fuzzy preference relations degenerate to crisp preference relations if $k=+\infty$. In this context, $\underline{R^{>}}d_i^{\geqslant}(x) = 1$ if all the samples better than x are in the preferential decision class $d_i^{\geqslant}$, and vice versa. According to the crisp preference based rough sets, $x \in \underline{R^{\geqslant}}d_i^{\geqslant}$ if $[x]_B^{\geqslant} \subseteq d_i^{\geqslant}$.

Proposition 1. *Given a fuzzy preference decision table* $< U,C,D >, a \subseteq C, x,y \in U$. *The value domain of* D *is* $\{d_i, i=1,\dots,N\}$. *We assume* $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_N$. $R^{>}$ *and* $R^{<}$ *are fuzzy preference relations generated by criterion a and functions* $\frac{1}{1+e^{-k(x_i-x_j)}}$ *and* $\frac{1}{1+e^{k(x_i-x_j)}}$. *If* $a(x) \geqslant a(y), \underline{R^{>}}d_i^{\geqslant}(x) \geqslant \underline{R^{>}}d_i^{\geqslant}(y), \underline{R^{<}}d_i^{\leqslant}(x) \leqslant \underline{R^{<}}d_i^{\leqslant}(y), \forall i \in \{1,2,\dots,N\}$.

Proof. Assuming that $R^{<}(z,x) = \inf_{u \notin d_i^{\geqslant}} R^{<}(u,x)$, we have $\underline{R^{>}} d_i^{\geqslant}(x) = R^{<}(z,x)$. It is easy to see that $\underline{R^{>}} d_i^{\geqslant}(y) = R^{<}(z,y)$. By $a(x) \geqslant a(y), R^{<}(z,x) \geqslant R^{<}(z,y)$ according to transitivity of fuzzy preference relations. Finally, we obtain the conclusion that $\underline{R^{>}} d_i^{\geqslant}(x) \geqslant \underline{R^{>}} d_i^{\geqslant}(y)$ if $a(x) \geqslant a(y)$. Similarly, $\underline{R^{<}} d_i^{\leqslant}(x) \leqslant \underline{R^{<}} d_i^{\leqslant}(y)$. □

Proposition 2. *Given a fuzzy preference decision table $\langle U, C, D \rangle, x \in U, B_1, B_2 \subseteq C, R_1^{>}, R_1^{<}$ are fuzzy preference relations generated by B_1; $R_2^{>}, R_2^{<}$ are fuzzy preference relations generated by B_2. Then the upwards and downwards fuzzy preference relations generated by $B_1 \bigcup B_2$ are $R^{>} = \min(R_1^{>}, R_2^{>})$ and $R^{<} = \min(R_1^{<}, R_2^{<})$, respectively. For $\forall x \in d_i^{\geqslant}, y \in d_i^{\leqslant}$, we have $\underline{R^{>}} d_i^{\geqslant}(x) \geqslant \underline{R_1^{>}} d_i^{\geqslant}(x)$, $\underline{R^{>}} d_i^{\geqslant}(x) \geqslant \underline{R_2^{>}} d_i^{\geqslant}(x)$ and $\underline{R^{<}} d_i^{\leqslant}(y) \geqslant \underline{R_1^{>}} d_i^{\leqslant}(y), \underline{R^{<}} d_i^{\leqslant}(y) \geqslant \underline{R_2^{<}} d_i^{\leqslant}(y)$.*

Proof. According to the definitions of fuzzy preference based rough sets, $\underline{R_1^{>}} d_i^{\geqslant}(x) = \inf_{u \in U} \max(1 - R_1^{>}(u,x), d_i^{\geqslant}(u))$; $\underline{R_2^{>}} d_i^{\geqslant}(x) = \inf_{u \in U} \max(1 - R_2^{>}(u,x), d_i^{\geqslant}(u))$; $\underline{R^{>}} d_i^{\geqslant}(x) = \inf_{u \in U} \max(1 - R^{>}(u,x), d_i^{\geqslant}(u))$. By $R^{>} = \min(R_1^{>}, R_2^{>})$, we obtain $1 - R^{>}(u,x) \geqslant 1 - R_1^{>}(u,x)$ and $1 - R^{>}(u,x) \geqslant 1 - R_2^{>}(u,x)$. Thus $\max(1 - R^{>}(u,x), d_i^{\geqslant}(u)) \geqslant \max(1 - R_1^{>}(u,x), d_i^{\geqslant}(u))$ and $\max(1 - R^{>}(u,x), d_i^{\geqslant}(u)) \geqslant \max(1 - R_2^{>}(u,x), d_i^{\geqslant}(u))$. Finally we derive that $\inf_{u \in U} \max(1 - R^{>}(u,x), d_i^{\geqslant}(u)) \geqslant \inf_{u \in U} \max(1 - R_1^{>}(u,x), d_i^{\geqslant}(u))$ and $\inf_{u \in U} \max(1 - R^{>}(u,x), d_i^{\geqslant}(u)) \geqslant \inf_{u \in U} \max(1 - R_2^{>}(u,x), d_i^{\geqslant}(u))$. Similarly, we can also see that $\underline{R^{<}} d_i^{\leqslant}(y) \geqslant \underline{R_1^{>}} d_i^{\leqslant}(y), \underline{R^{<}} d_i^{\leqslant}(y) \geqslant \underline{R_2^{<}} d_i^{\leqslant}(y)$. □

There are multiple criteria in a preference decision table. The more criteria we can use, the more the decision is consistent. Proposition 1 indicates that adding new criteria into the current subset, the certainty of decision improves.

Definition 4. Given $\langle U, C, D \rangle, R^{>}$ and $R^{<}$ are two fuzzy preference relations generated by $B \subseteq C$. For $\forall d_i \in D$, we introduce a coefficient, called the fuzzy preference approximation quality (FPAQ), to characterize the qualities of approximations:

- upward FPAQ: $\gamma_B^{>}(d_i^{\geqslant}) = \frac{\sum_{x \in d_i^{\geqslant}} \underline{R^{>}} d_i^{\geqslant}(x)}{|d_i^{\geqslant}|}$,
- downward FPAQ: $\gamma_B^{<}(d_i^{\leqslant}) = \frac{\sum_{x \in d_i^{\leqslant}} \underline{R^{<}} d_i^{\leqslant}(x)}{|d_i^{\leqslant}|}$,
- global FPAQ: $\gamma_B(d_i) = \frac{\sum_{x \in d_i^{\leqslant}} \underline{R^{<}} d_i^{\leqslant}(x) + \sum_{x \in d_i^{\geqslant}} \underline{R^{>}} d_i^{\geqslant}(x)}{|d_i^{\geqslant}| + |d_i^{\leqslant}|}$,

where $|d_i^{\geqslant}|$ and $|d_i^{\leqslant}|$ are the numbers of samples with decisions dominating and dominated by d_i, respectively. Since $0 \leqslant \underline{R^{>}} d_i^{\geqslant}(x) \leqslant 1$ and $0 \leqslant \underline{R^{<}} d_i^{\leqslant}(x) \leqslant 1$, we have $0 \leqslant \gamma_B^{>}(d_i^{\geqslant}) \leqslant 1, 0 \leqslant \gamma_B^{<}(d_i^{\leqslant}) \leqslant 1$ and $0 \leqslant \gamma_B(d_i) \leqslant 1$. We say d_i is upward, downward or global definable if $\gamma_B^{>}(d_i^{\geqslant}) = 1, \gamma_B^{<}(d_i^{\leqslant}) = 1$ and $\gamma_B(d_i) = 1$, respectively.

Definition 5. Given $\langle U, C, D \rangle, R^{>}$ and $R^{<}$ are two fuzzy preference relations generated by $B \subseteq C$. The value domain of D is $\{d_i, i = 1, \ldots, N\}$. We assume $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_N$. The fuzzy preference approximation qualities of D with respect to B are defined as:

- upward FPAQ: $\gamma_B^{>}(D^{\geqslant}) = \frac{\sum_i \sum_{x \in d_i^{\geqslant}} \underline{R^{>}} d_i^{\geqslant}(x)}{\sum_i |d_i^{\geqslant}|}$,
- downward FPAQ: $\gamma_B^{<}(D^{\leqslant}) = \frac{\sum_i \sum_{x \in d_i^{\leqslant}} \underline{R^{<}} d_i^{\leqslant}(x)}{\sum_i |d_i^{\leqslant}|}$,
- global FPAQ: $\gamma_B(D) = \frac{\sum_i \left(\sum_{x \in d_i^{\leqslant}} \underline{R^{<}} d_i^{\leqslant}(x) + \sum_{x \in d_i^{\geqslant}} \underline{R^{>}} d_i^{\geqslant}(x) \right)}{\sum_i \left(|d_i^{\geqslant}| + |d_i^{\leqslant}| \right)}$.

Obviously, $0 \leqslant \gamma_B^{>}(D^{\geqslant}) \leqslant 1, 0 \leqslant \gamma_B^{<}(D^{\leqslant}) \leqslant 1$ and $0 \leqslant \gamma_B(D) \leqslant 1$. We say that D is upward, downward or global consistent if $\gamma_B^{>}(D^{\geqslant}) = 1, \gamma_B^{<}(D^{\leqslant}) = 1$ or $\gamma_B(D) = 1$, respectively.

Proposition 3. *Given a globally consistent fuzzy preference decision table $< U, C, D >, R^{>}$ and $R^{<}$ are fuzzy preference relations induced by C. We have $\gamma_C^{>}(D^{\geqslant}) = 1, \gamma_C^{<}(D^{\leqslant}) = 1$, and for $\forall x \in d_i^{\geqslant}, \underline{R^{>}} d_i^{\geqslant}(x) = 1$ and for $\forall y \in d_i^{\leqslant}, \underline{R^{<}} d_i^{\leqslant}(y) = 1$.*

Proof. If $\langle U, C, D \rangle$ is consistent, $\gamma_C^{>}(D^{\geqslant}) = 1, \gamma_C^{<}(D^{\leqslant}) = 1$. According to definition, we know $\gamma_C^{>}(D^{\geqslant}) = \frac{\sum_i \sum_{y \in d_i^{\geqslant}} \underline{R^{>}} d_i^{\geqslant}(y)}{\sum_i |d_i^{\geqslant}|}$ and $\gamma_C^{<}(D^{\leqslant}) = \frac{\sum_i \sum_{y \in d_i^{\leqslant}} \underline{R^{<}} d_i^{\leqslant}(y)}{\sum_i |d_i^{\leqslant}|}$. As shown before $0 \leqslant \underline{R^{<}} d_i^{\leqslant}(y) \leqslant 1$, we know that $\sum_{y \in d_i^{\leqslant}} \underline{R^{<}} d_i^{\leqslant}(y) < |d_i^{\leqslant}|$ if $\exists y \in d_i^{\leqslant}, \underline{R^{<}} d_i^{\leqslant}(y) < 1$, and then $\frac{\sum_{y \in d_i^{\leqslant}} \underline{R^{<}} d_i^{\leqslant}(y)}{|d_i^{\leqslant}|} < 1$. Therefore, $\gamma_C^{<}(D^{\leqslant}) = \frac{\sum_i \sum_{y \in d_i^{\leqslant}} \underline{R^{<}} d_i^{\leqslant}(y)}{\sum_i |d_i^{\leqslant}|} < 1$. So $\forall x \in d_i^{\leqslant}$, we have $\underline{R^{<}} d_i^{\leqslant}(y) = 1$. Analogically, we can see that $\forall x \in d_i^{\geqslant}, \underline{R^{>}} d_i^{\geqslant}(x) = 1$. □

Proposition 4. *Given a fuzzy preference decision table $\langle U, C, D \rangle, a \in B \subseteq C$. We have $\gamma_B^{>}(D^{\geqslant}) \geqslant \gamma_{B-\{a\}}^{>}(D^{\geqslant}), \gamma_B^{<}(D^{\leqslant}) \geqslant \gamma_{B-\{a\}}^{<}(D^{\leqslant})$ and $\gamma_B(D) \geqslant \gamma_{B-\{a\}}(D)$.*

Proof. We can easily derive the conclusion according to Proposition 2 and Definition 5. □

Definition 6. Given $\langle U,C,D\rangle, R^{>}$ and $R^{<}$ are two fuzzy preference relations generated by $B \subseteq C$. The value domain of D is $\{d_i, i=1,\ldots,N\}$, and $d_1 \leqslant d_2 \leqslant \cdots \leqslant d_N$. $a \in B, S^{>}$ and $S^{<}$ are two fuzzy preference relations generated by $B-a$. We say a is upward redundant if $\gamma_B^{\geqslant}(D^{\geqslant}) = \gamma_{B-\{a\}}^{\geqslant}(D^{\geqslant})$; otherwise a is upward indispensable and $\gamma_B^{\geqslant}(D^{\geqslant}) > \gamma_{B-\{a\}}^{\geqslant}(D^{\geqslant})$. Similarly, we say a is downward redundant, globally redundant, downward indispensable or globally indispensable if $\gamma_B^{\leqslant}(D^{\leqslant}) = \gamma_{B-\{a\}}^{\leqslant}(D^{\leqslant}), \gamma_B(D) = \gamma_{B-\{a\}}(D), \gamma_B^{\leqslant}(D^{\leqslant}) > \gamma_{B-\{a\}}^{\leqslant}(D^{\leqslant}), \gamma_B(D) > \gamma_{B-\{a\}}(D)$.

Definition 7. Given a fuzzy preference decision table $\langle U,C,D\rangle, B \subseteq C$, we say B is an upward reduct if

(1) $\gamma_B^{\geqslant}(D^{\geqslant}) = \gamma_C^{\geqslant}(D^{\geqslant})$;
(2) for $\forall a \in B \gamma_B^{\geqslant}(D^{\geqslant}) > \gamma_{B-\{a\}}^{\geqslant}(D^{\geqslant})$.

Correspondingly, we can define the downward reducts and global reducts as follows.

Definition 8. Given a fuzzy preference decision table $\langle U,C,D\rangle, B \subseteq C$, we say B is an downward reduct if

(1) $\gamma_B^{\leqslant}(D^{\leqslant}) = \gamma_C^{\leqslant}(D^{\leqslant})$;
(2) for $\forall a \in B \gamma_B^{\leqslant}(D^{\leqslant}) > \gamma_{B-\{a\}}^{\leqslant}(D^{\leqslant})$.

Definition 9. Given a fuzzy preference decision table $\langle U,C,D\rangle, B \subseteq C$ we say B is an global reduct if

(1) $\gamma_B(D) = \gamma_C(D)$;
(2) for $\forall a \in B \gamma_B(D) > \gamma_{B-\{a\}}(D)$.

5. Preference analysis with fuzzy preference approximation

Given a preference decision table, we usually require an analysis of the relevance between criteria and decisions. We want to know which criteria are significant for decision-making and which criteria are redundant or irrelevant to the decision. Thus we can obtain an indispensable and sufficient subset of criteria for constructing preference decision models. In this case, we introduce coefficients and algorithms for analyzing fuzzy preference decision-making based on fuzzy preference approximation.

In the above section, we introduce the measure of approximation quality. These definitions can be used to evaluate the approximation capability of a criterion or a set of criteria.

Definition 10. Given $\langle U,C,D\rangle, R^{>}$ and $R^{<}$ are two fuzzy preference relations generated by $B \subseteq C$. The capability of B to approximate D are defined as:

- Upward approximating capability (UAC): $\gamma_B^{\geqslant}(D^{\geqslant}) = \frac{\sum_i \sum_{x \in d_i^{\geqslant}} \underline{R^{>}d_i^{\geqslant}}(x)}{\sum_i |d_i^{\geqslant}|}$,
- downward approximating capability (DAC): $\gamma_B^{\leqslant}(D^{\leqslant}) = \frac{\sum_i \sum_{x \in d_i^{\leqslant}} \underline{R^{<}d_i^{\leqslant}}(x)}{\sum_i |d_i^{\leqslant}|}$,
- global approximating capability (GAC): $\gamma_B(D) = \frac{\sum_i \left(\sum_{x \in d_i^{\leqslant}} \underline{R^{<}d_i^{\leqslant}}(x) + \sum_{x \in d_i^{\geqslant}} \underline{R^{>}d_i^{\geqslant}}(x) \right)}{\sum_i \left(|d_i^{\geqslant}| + |d_i^{\leqslant}| \right)}$.

We can compute the dependency or relevance between criteria and decisions with the above definitions. We can also compute the significance of a single criterion when a set of criteria have been given.

Definition 11. Given $\langle U,C,D\rangle, R^{>}$ and $R^{<}$ are two fuzzy preference relations generated by $B \subseteq C$ and $S^{>}$and $S^{<}$ are two fuzzy preference relations generated by $B \bigcup a$. The conditional significances of a relative to B to approximate D are defined as:

- Upward conditional significance:

$$sig^{>}(a,B,D^{\geqslant}) = \frac{\sum_i \sum_{x \in d_i^{\geqslant}} \left(\underline{S^{>}d_i^{\geqslant}}(x) - \underline{R^{>}d_i^{\geqslant}}(x) \right)}{\sum_i |d_i^{\geqslant}|},$$

- downward conditional significance:

$$sig^{<}(a,B,D^{\leqslant}) = \frac{\sum_i \sum_{x \in d_i^{\leqslant}} \left(\underline{S^{<}d_i^{\leqslant}}(x) - \underline{R^{<}d_i^{\leqslant}}(x) \right)}{\sum_i |d_i^{\leqslant}|},$$

- global conditional significance:

$$sig(a,B,D)=\frac{\sum_i\left(\sum_{x\in d_i^{\leqslant}}\left(\underline{S^{<}}d_i^{\leqslant}(x)-\underline{R^{<}}d_i^{\leqslant}(x)\right)+\sum_{x\in d_i^{\geqslant}}\left(\underline{S^{>}}d_i^{\geqslant}(x)-\underline{R^{>}}d_i^{\geqslant}(x)\right)\right)}{\sum_i\left(|d_i^{\geqslant}|+|d_i^{\leqslant}|\right)}.$$

Attribute reduction and feature selection are the key preprocessing steps for machine learning and pattern recognition. Selecting the relevant and informative criteria also play an important role in decision-making modeling. Based on the above definitions of criterion significance and some search strategy, we can develop an algorithm for criteria selection in fuzzy preference based rough sets. Considering efficiency of computation, we introduce a forward greedy search algorithm to find a reduct of a preference decision table.

Algorithm 1. Forward greedy search of a reduct based on fuzzy preference rough sets.

Input: preference decision table $\langle U,C,D\rangle$
Output: an ordinal reduct of the decision table
Step 1: $\emptyset\to red$; // red is the pool to contain the selected attributes
Step 2: For each $a_i\in A-red$
  compute $sig(a_i,B,D)$, // Here we define $\gamma_{\emptyset}(D)=0$
 end
Step 3: select the attribute a_k which satisfies:

$$sig(a_k,red,D)=\max_i(sig(a_i,red,D))$$

Step 4: if $sig(a_k,red,D)>0$,

$$red\bigcup a_k\to red$$

  go to step2
 else
  return red
Step 5: end

In Algorithm 1, we can select different definitions of attribute significance, such as $sig^{>}(a,red,D^{\geqslant}),sig^{<}(a,red,D^{\leqslant})$ or $sig(a,red,D)$, then we will get an upwards reduct, a downwards reduct or a global reduct, respectively. We show an algorithm which computes the upwards conditional significance of the attributes in Algorithm 2. The other two measures of attribute significance can be constructed in a similar way.

Algorithm 2. Compute the upward conditional significance of attribute a

Input: $\langle U,C,D\rangle,B\subseteq C,a\in C$ and $a\notin B$;
Output: $sig^{>}(a,B,D)$;
Step 1: compute $\sum_i|d_i^{\geqslant}|$
Step 2: compute the fuzzy preference relations induced by B and $B\bigcup\{a\}$, denoted by $R^{>}$and $S^{>}$
Step 3: $r_1\leftarrow 0,r_2\leftarrow 0$
Step 4: for $i=1$ to n
  find the ordinal number of the decision of x, assume it is k
  for $j=1$ to k
  compute $LA_1(j)=\underline{R^{>}}d_j^{\geqslant}(x_i),LA_2(j)=\underline{S^{>}}d_j^{\geqslant}(x_i)$
  end

$$LA_1(i)\leftarrow\sum_i LA_1(j),LA_2(i)\leftarrow\sum_i LA_2(j)$$

  end
step 5: $r_1\leftarrow\sum_i LA_1(i),r_2\leftarrow\sum_i LA_2(i)$
step 6: $sig^{>}(a,B,D)\leftarrow(r_2-r_1)/\sum_i|d_i^{\geqslant}|$,
step 7. end

To compute the significance of attributes, the fuzzy preference relation between samples should be computed, which is time-consuming if all the sample pairs are calculated. We can rank the samples according to the attribute values before we compute relations, so that we can compare the samples in a small domain. For example, with respect to a sample x, we define a computational domain as $[x-\varepsilon, x+\varepsilon]$. If $y \in [x-\varepsilon, x+\varepsilon], R^{>}(x,y) = \frac{1}{1+e^{-k(x-y)}}$. If $y < x-\varepsilon, R^{>}(x,y) = 1$ and if $y > x-\varepsilon, R^{>}(x,y) = 0$. In this way, we need to only compute the relations between x and the samples in the computational domain. Assuming there are n samples, among which k samples are located in the computational domain of each sample, then the time complexity in computing significance of an attribute is $O(n\log n + kn)$, where $n\log n$ is for ranking and kn is for computing fuzzy preference. Moreover, if there are N candidate features in the decision table, the time complexity in searching reduct is $O((n\log n + kn)N^2)$.

6. Experiments

Consider the problem of computing the dependency of decisions on attribute originality and writing in Table 1. We try both the dominance rough set model and the proposed fuzzy preference based rough set model, respectively. The derived results are shown in Table 4. We change the attribute originality of sample x_3 from 0.6 to 0.74. Intuitively, the inconsistency of a decision based on the originality should become higher in this case. We compute the dependency of the originality and the writing quality based on dominance rough sets and fuzzy preference based rough sets again. The results are shown in Table 5. Although the inconsistency becomes greater, there is no difference between the dependencies computed with

Table 4
Dependency computed with Example 1.

	Crisp preference rough set		Fuzzy preference rough set	
	Originality	Writing	Originality	Writing
Upward	0.6000	1.0000	0.7988	0.9159
Downward	0.8000	1.0000	0.8900	0.9362
Global	0.7000	1.0000	0.8444	0.9261

Table 5
Dependency after a sample is revised.

	Crisp preference rough set		Fuzzy preference rough set	
	Originality	Writing	Originality	Writing
Upward	0.6000	1.0000	0.7270	0.9159
Downward	0.8000	1.0000	0.8702	0.9362
Global	0.7000	1.0000	0.7986	0.9261

Table 6
Attribute of pasture production.

ID	Name	Description	
1.	Fertiliser	Fertiliser used {LL, LN, HN, HH}	Enumerated
2.	Slope	Slope of the paddock	Integer
3.	Aspect-dev-NW	The deviation from the north-west	Integer
4.	OlsenP		Integer
5.	MinN		Integer
6.	TS		Integer
7.	Ca-Mg	Calcium magnesium ration	Real
8.	LOM	Soil lom (g/100 g)	Real
9.	NFIX-mean	A mean calculation	Real
10.	Eworms-main-3	Main 3 spp earth worms per g/m^2	Real
11.	Eworms-No-species	Number of spp	Integer
12.	KUnSat	mm/h	Real
13.	OM		Real
14.	Air-Perm		Real
15.	Porosity		Real
16.	HFRG-pct-mean	Mean percent	Real
17.	Legume-yield	kgDM/ha	Real
18.	OSPP-pct-mean	Mean percent	Real
19.	Jan-Mar-mean-TDR		Real
20.	Annual-Mean-Runoff	In mm	Real
21.	Root-surface-area	m^2/m^3	Real
22.	Leaf-P	ppm	Real
Class	Pasture prod-class	Pasture production categorization {Low, Median, High}	Enumerated

dominance rough sets. As to fuzzy preference rough sets based dependency, the values of dependency become smaller after the revision. This experiment shows that dominance rough sets can not reflect the degree of preference, while fuzzy preference based rough sets are able to measure this kind of information.

Two data sets, named Pasture Production and Squash Harvest, were downloaded from the web site (http://www.cs.waikato.ac.nz/ml/weka/). Pasture Production was collected by Dave Barker. The objective was to predict pasture production from a variety of biophysical factors. Vegetation and soil variables from areas of grazed North Island hill country with different management (fertilizer application/stocking rate) histories (1973–1994) were measured and subdivided into 36 paddocks. Nineteen vegetation (including herbage production); soil chemical, physical and biological; and soil water variables were selected as potentially useful. There are 36 instances characterized by 22 attributes in the database. Pasture production categorization is {LO, MED, HI}. The name and the corresponding description of these attributes are described in Table 6.

There are 15 real-valued, six integer-valued and one enumerated criteria in the data. These attributes are not of the same importance in predicting pasture productions. The estimation of relevance of these criteria to the pasture production is useful for improving pasture cultivation. Before computing the dependency between the attribute and decisions, we should know the monotonous information between attribute and decision. In application, users usually know this relationship with

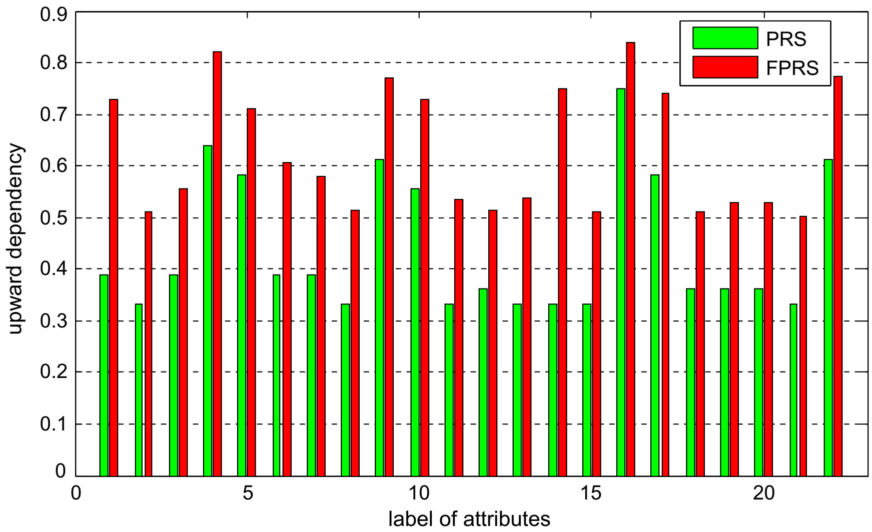

Fig. 3. Upward dependency based on preference rough sets (PRS) and fuzzy preference rough sets (FPRS).

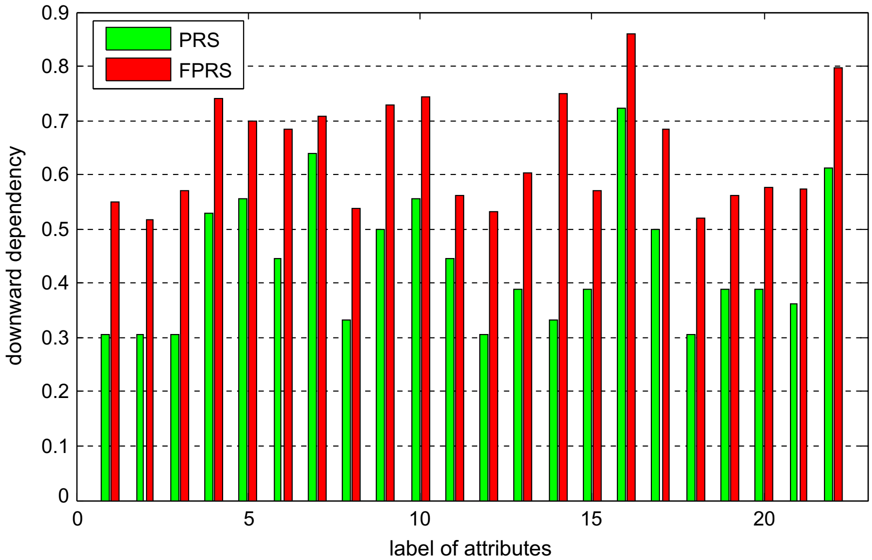

Fig. 4. Downward dependency based on preference rough sets (PRS) and fuzzy preference rough sets (FPRS).

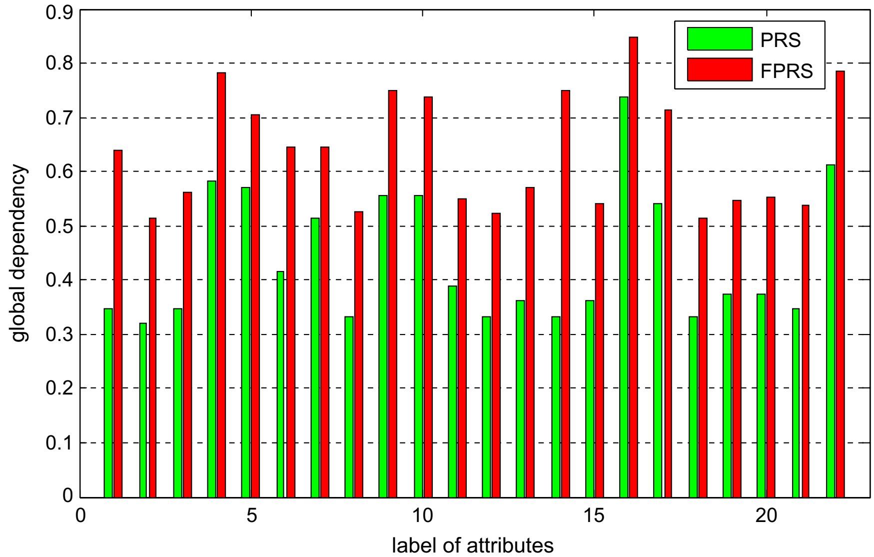

Fig. 5. Global dependency based on preference rough sets (PRS) and fuzzy preference rough sets (FPRS).

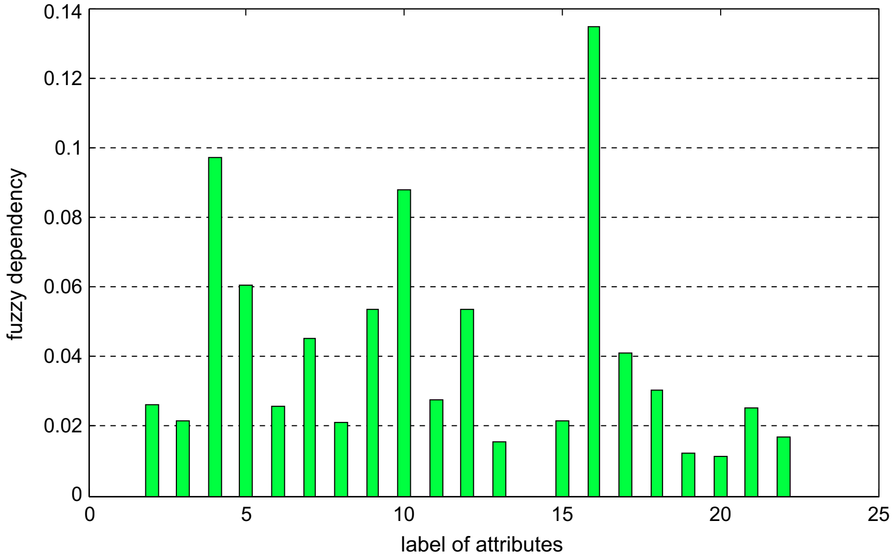

Fig. 6. Fuzzy dependency based on fuzzy rough sets.

their domain knowledge. Here we do not discuss how to get the monotonicity. And we use the assumption that if a sample x contains greater values than another sample y in terms of any attribute, the decision of x should not be worse than y. For example, we assume that the greater the slope is, the higher the pasture production is. Based on this assumption, we compute the dependency of production on each single criterion based on crisp preference rough sets and fuzzy preference based rough sets. Figs. 3–5 show the comparison of upward dependency, downward dependency and global dependency of a single attribute, while Fig. 6 presents dependency of each single attribute computed with fuzzy rough sets. The information of ordinal structure in decision is neglected in the fuzzy rough set model.

As a whole, there is a great difference between the fuzzy dependency computed with fuzzy rough sets and other measures. This result is reasonable because the fuzzy rough set model does not consider the ordinal information while other models reflect the ordinal structure in decision classes. Moreover, in upward dependency computed by preference rough sets, we sort the attributes in a descending order {16, 4, 9, 22, 5, 17, 10, 1, 3, 6, 7, 12, 18, 19, 20, 2, 8, 11, 13, 14, 15, 21}. When upward dependency is computed with fuzzy preference based rough sets, the order is {16, 4, 22, 9, 14, 17, 10, 1, 5, 6, 7, 3, 13, 11,

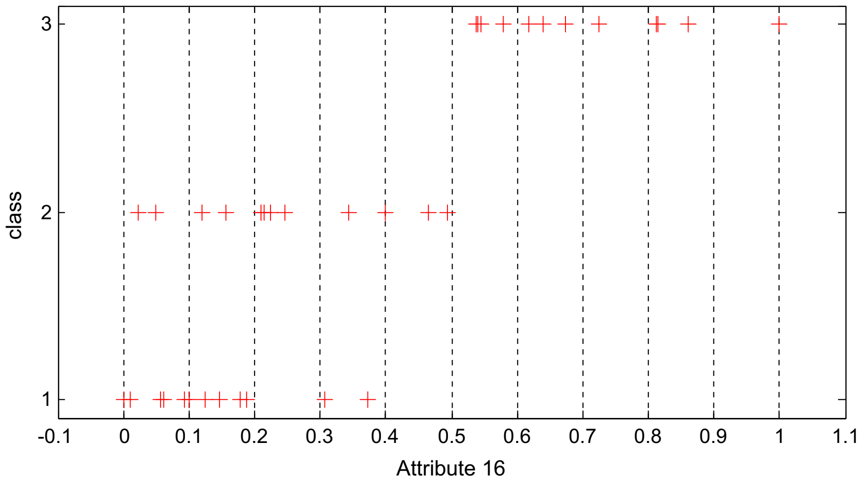

Fig. 7. Scattering plot of attribute 16 over decision.

Table 7
Reduct and dependency computed with different evaluation methods.

Method	Reduct	Dependency of attribute subset
Upward PRS	16, 4, 11	0.7500, 0.9444, 1
Upward FPRS	16, 4, 10, 11	0.8394, 0.9274, 0.9463, 0.9610
Downward PRS	16, 10, 4	0.7222, 0.9167, 0.9722
Downward FPRS	16, 10, 5, 22	0.8586, 0.9141, 0.9454, 0.9620
Global PRS	16, 17, 13	0.7361, 0.9167, 0.9861
Global FPRS	16, 4, 11, 8, 5	0.8490, 0.9126, 0.9411, 0.9529, 0.9652
Fuzzy rough sets	16, 1, 11, 13, 15, 8	0.1348, 0.4400, 0.7960, 0.9079, 0.9734, 0.9929

Table 8
Attributes of squash harvest.

ID	Name	Description	Value class
1.	Site	Where fruit is located {P,HB,LINC}	Enumerated
2.	Daf	Number of days after flowering {30, 40, 50, 60, 70}	Enumerated
3.	Druit	Individual number of the fruit (not unique)	Enumerated
4.	Weight	Weight of whole fruit in grams	real
5.	Storewt	Weight of fruit after storage	real
6.	Pene	Penetrometer indicates maturity of fruit at harvest	Integer
7.	Solids_%	A test for dry matter	Integer
8.	Brix	A refractometer measurement used to indicate sweetness or ripeness of the fruit	Integer
9.	a^*	The a^* coordinate of the HunterLab L^* a^* b^* notation of colour measurement	Integer
10.	egdd	The heat accumulation above a base of 8c from emergence of the plant to harvest of the fruit	real
11.	fgdd	The heat accumulation above a base of 8c from flowering to harvesting	Real
12.	Groundspot_a^*	The number indicating colour of skin where the fruit rested on the ground	Integer
13.	Glucose	Measured in mg/100 lg of fresh weight	Integer
14.	Fructose	Measured in mg/100 g of fresh weight	Integer
15.	Sucrose	Measured in mg/100g of fresh weight	Integer
16.	Total	Measured in mg/100 g of fresh weight	Integer
17.	Glucose + fructos	Measured in mg/100 g of fresh weight	Integer
18.	Starch	Measured in mg/100 g of fresh weight	Integer
19.	Sweetness	The mean of eight taste panel scores; out of 1500	Integer
20.	Flavour	The mean of eight taste panel scores; out of 1500	Integer
21.	Dry/moist	The mean of eight taste panel scores; out of 1500	Integer
22.	Fibre	The mean of eight taste panel scores; out of 1500	Integer
23.	Heat_input_emerg	The amount of heat emergence after harvest	real
24.	Heat_input_flower	The amount of heat input before flowering	real
25.	Acceptability	The acceptability of the fruit {excellent, ok, not_acceptable}	Enumerated

20, 19, 12, 8, 15, 2, 18, 21}. The order of Features 5, 14 and some others significantly vary with respect to upward dependency. A similar case occurs for downward dependency and global dependency.

We also find that some attributes produce greater or smaller dependency values for all kinds of computations of dependency, such as attribute 16 and attribute 21. We generate the scatter plot of Feature 16 over decisions, as shown in Fig. 7. The samples with feature values greater than 0.5 all belong to class 3. Therefore, these samples are consistent with respect to all computations mentioned above. Accordingly, we can come to the conclusion that pasture production is high if the HFRG-pct-mean is greater than 0.5.

Finally, we compute the reduct of criteria based on different computations of dependency. The results are shown in Table 7, where dependency is computed with the selected features. For example, dependency values of 0.7500, 0.9444 and 1 are computed with feature subsets {16}, {16,4} and {16,4,11}, respectively. We can see that attributes 16, 4, 11, 10 appear in most of the reducts. This indicates that these features are important with respect to different evaluation functions.

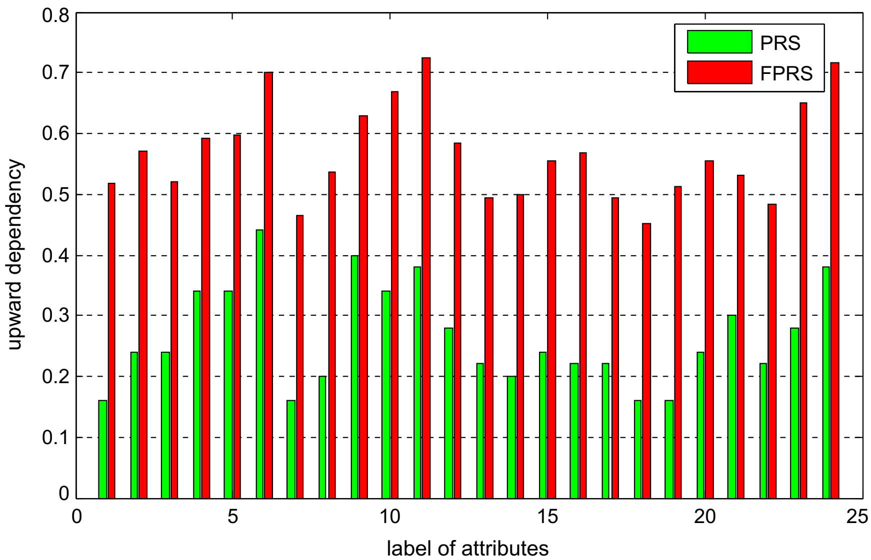

Fig. 8. Upward dependency based on preference rough sets (PRS) and fuzzy preference rough sets (FPRS).

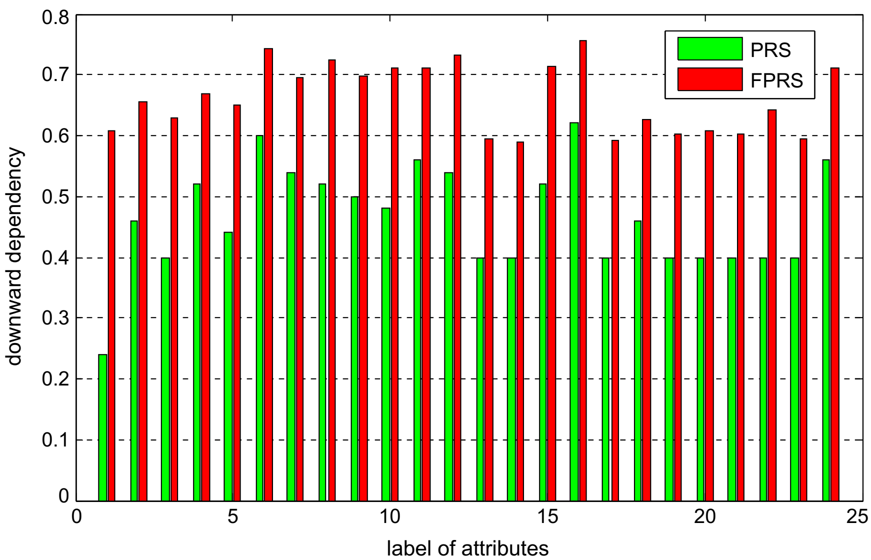

Fig. 9. Downward dependency based on preference rough sets (PRS) and fuzzy preference rough sets (FPRS).

The second data set is the squash harvest, donated by Winna Harvey in Crop and Food Research, Christchurch, New Zealand. The purpose of the research was to determine the changes taking place in squash fruit during the maturation and ripening so as to pinpoint the best time to give the best quality in the market place (Japan). The squash is transported to Japan by refrigerated cargo vessels and takes three to four weeks to reach the market. Evaluations were carried out at two stages: prior to export and after arriving at the market.

The original objectives were to determine which pre-harvest variables contribute to good tasting squash after different periods of storage time. This is determined by categorizing each squash as either unacceptable, acceptable or excellent. The numbers of instances and attribute are 52 and 24, respectively. The attributes and their descriptions are listed in Table 8.

Similarly, we also give the upward dependency, downward dependency and global dependency computed with preference rough sets and fuzzy preference rough sets, as shown in Figs. 8–10. For upward and global dependency computed with preference rough sets, Feature 6 produces the maximal dependency compared to the other features. For downward dependency, whether crisp or fuzzy, Feature 16 generates the greatest dependency among 24 features. As a whole, considering both upward consistency and downward consistency, the taste of squash largely depends on maturity of fruit at harvest. The computational result is consistent with our common knowledge that the taste of squash is relevant to maturity.

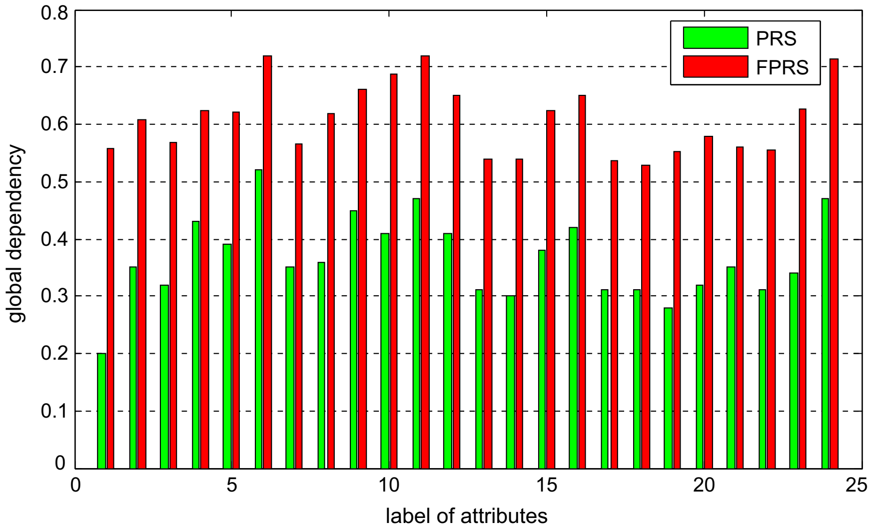

Fig. 10. Global dependency based on preference rough sets (PRS) and fuzzy preference rough sets (FPRS).

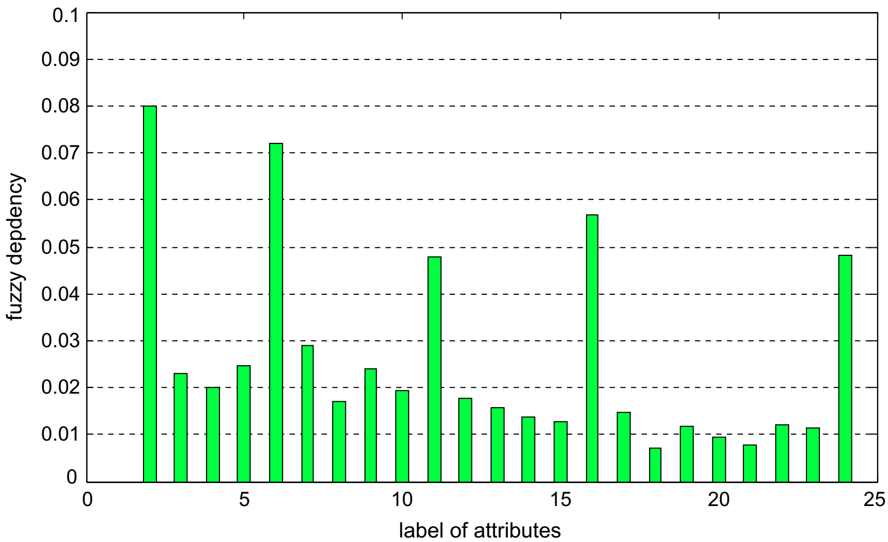

Fig. 11. Dependency based on fuzzy rough sets (FRS).

Table 9
Reduct and dependency computed with different evaluation methods.

Method	Reduct	Dependency of attribute subset
Upward PRS	6, 22, 4, 19, 1	0.4400, 0.8600, 0.9400, 0.9800, 1.0000
Upward FPRS	11, 14, 15, 22, 2, 4	0.7242, 0.8202, 0.8884, 0.9117, 0.9314, 0.9449
Downward PRS	16, 7, 20	0.6200, 0.7600, 0.8200
Downward FPRS	16, 11, 14, 7, 4, 2, 21	0.7550, 0.8516, 0.8968, 0.9267, 0.9382, 0.9500, 0.9602
sglobal PRS	6, 22, 4, 19	0.5200, 0.7900, 0.8500, 0.8700
Global FPRS	6, 14, 11, 2, 4, 21, 18	0.7200, 0.8500, 0.8935, 0.9204, 0.9376, 0.9502, 0.9632
Fuzzy rough sets	2, 1, 3, 15, 13, 21, 18, 12, 19	0.0800, 0.3000, 0.5728, 0.7599, 0.8563, 0.9065, 0.9383, 0.9601, 0.9741

It is notable that Feature 1 outputs dependency zero when we compute dependency with fuzzy rough sets, as shown in Fig. 11. It is greater than zero if other computations are used. Moreover, according to Feature 2, the number of days after flowering gets the maximal dependency in fuzzy rough sets. The conclusions are not reasonable since ordinal information is overlooked in fuzzy rough sets where fuzzy similarity relations, rather than fuzzy preference relations, are used.

The reducts and dependency of attribute subsets are given in Table 9. Features 6, 22, 4, 19, 1 are required for getting consistent upward decisions in terms of crisp preference based rough sets. Features 11, 14, 15, 22, 2 and 4 are selected by fuzzy preference based rough sets. The difference comes from the definitions of dependency. In fuzzy preference based rough sets, the membership of a sample to the lower approximation of its preference decision is small even though it is consistent when there is a sample close to it that has a different preference decision.

7. Conclusions and future work

Preference analysis is an important task in intelligent data analysis and machine learning. The classical rough set theory was successfully extended to dealing with preference analysis by replacing equivalence relations with dominance relations. Dominance relations cannot reflect the fuzziness presented in criteria. In this paper, we introduce the logsig function, which is widely used in BP neural networks, to extract fuzzy preference relations from samples. We integrate fuzzy preference relations with an improved fuzzy rough set model, and develop a fuzzy preference based rough set model. The relations between the proposed model and the crisp preference based rough set model are discussed. Furthermore, we generalize the dependency used in classical rough sets and fuzzy rough sets to compute the relevance between the criteria and decisions. We propose the definitions of upward dependency, downward dependency and global dependency.

We use the paper reviewing decision to explain the model and obtain an interesting conclusion that the lower and upper approximations in fuzzy preference based rough sets can be understood as the pessimistic and optimistic decisions in human reasoning, respectively. A pessimistic decision maker thinks his paper would be accepted with greater possibility if it is better than the best paper which is rejected; while an optimistic decision maker thinks that his paper would be accepted with greater possibility if it is better than the worst paper accepted. The fuzzy lower approximation operator computes how much this paper is better than the best rejected one, while the fuzzy upper approximation operator measures how much this paper is better than the worst accepted one.

The proposed model is used to analyze two publicly available data sets. One is about pasture production, and the other is about the taste of squash. We compute the dependency of decision on the available criteria, which reflects the relevance between criteria and decisions. The derived results are consistent with human intuition, which shows the effectiveness of the proposed model.

There are some problems to be addressed in fuzzy preference based approximations. Firstly, although we introduce a function to compute the preference degrees of objects, there might be a number of functions to measure preferences. We should discuss them and the choices of their parameters. Secondly, it is an interesting issue to determine whether there are monotonous relations between attributes and decisions and the kinds of such relations. The answer is usually given by users according to their prior knowledge. However, users may have difficulties in providing such information in applications and we require an algorithm to find the relations. Finally, there is a problem with the proposed model. Given a consistent sample, we can find that the membership of this sample to the upper approximation is less than that of the lower approximation, which is not consistent with the definition of fuzzy approximation. We require a new model to overcome this problem.

Acknowledgements

The authors thank Prof. Yiyu Yao, Prof. Witold Pedrycz and the anonymous referees for their constructive comments. This work is supported by National Natural Science Foundation of China under Grants 60703013, 10978011 and 60932008. Prof. Yu is supported by National Science Fund for Distinguished Young Scholars under Grant 50925625. Dr. Hu is supported by China Postdoctoral Science Foundation and The Hong Kong Polytechnic University (G-YX3B).

References

[1] R.B. Bhatt, M. Gopal, FRCT: fuzzy-rough classification trees, Pattern Analysis and Applications 11 (2008) 73–88.
[2] J. Braszczynski, S. Greco, R. Slowinski, Multi-criteria classification – a new scheme for application of dominance-based decision rules, European Journal of Operational Research 181 (2007) 1030–1044.
[3] D.G. Chen, C.Z. Wang, Q.H. Hu, A new approach to attribute reduction of consistent and inconsistent covering decision systems with covering rough sets, Information Sciences 177 (2007) 3500–3518.
[4] M.R. Chmielewski, J.W. Grzymala-Busse, Global discretization of continuous attributes as preprocessing for machine learning, International Journal of Approximate Reasoning 15 (1996) 319–331.
[5] T.-Q. Deng, Y.M. Chen, W.L. Xu, Q.H. Dai, A novel approach to fuzzy rough sets based on a fuzzy covering, Information Sciences 177 (2007) 2308–2326.
[6] D. Dubois, H. Fargier, H. Prade, Ordinal and probabilistic representations of acceptance, Journal of Artificial Intelligence Research 22 (2004) 23–56.
[7] D. Dubois, H. Prade, Twofold fuzzy sets and rough sets – some issues in knowledge representation, Fuzzy Sets and Systems 23 (1987) 3–18.
[8] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems 17 (1990) 191–209.
[9] I. Duntsch, G. Gediga, Uncertainty measures of rough set prediction, Artificial intelligence 106 (1998) 109–137.
[10] F. Fernandez-Riverola, F. Diaz, J.M. Corchado, Reducing the memory size of a fuzzy case-based reasoning system applying rough set techniques, IEEE Transactions on Systems Man and Cybernetics Part C – Applications and Reviews 37 (2007) 138–146.
[11] S. Greco, M. Inuiguchi, R. Slowinsk, Dominance-based rough set approach using possibility and necessity measures, in: J.J. Alpigini et al., (Eds.), RSCTC, LNAI 2475, pp. 85–92, 2002.
[12] S. Greco, B. Matarazzo, R. Slowinski, The use of rough sets and fuzzy sets in MCDM, in: T. Gal, T. Stewart, T. Hanne (Eds.), Advances in Multiple Criteria Decision Making, Kluwer Academic Publishers, Boston, 1999, pp. 14.1–14.59 (Chapter 14).
[13] S. Greco, B. Matarazzo, R. Slowinski, Rough approximation of a preference relation by dominance relations, European Journal of Operational Research 117 (1999) 63–83.
[14] S. Greco, B. Matarazzo, R. Slowinski, Fuzzy dominance-based rough set approach, in: F. Masulli, R. Parenti, G. Pasi (Eds.), Advances in Fuzzy Systems and Intelligent Technologies, Shaker Publishing, Maastricht, 2000, pp. 56–66.
[15] S. Greco, B. Matarazzo, R. Slowinski, A fuzzy extension of the rough set approach to multicriteria and multiattribute sorting, in: J. Fodor, B. De Baets, P. Perny (Eds.), Preferences and Decisions under Incomplete Information, Physica-Verlag, Heidelberg, 2000, pp. 131–154.
[16] S. Greco, B. Matarazzo, R. Slowinski, Rough sets theory for multi-criteria decision analysis, European Journal of operational research 129 (2001) 1–47.
[17] S. Greco, B. Matarazzo, R. Slowinski, Rough approximation by dominance relations, International Journal of Intelligent Systems 17 (2002) 153–171.
[18] S. Greco, M. Inuiguchi, R. Slowinski, Dominance-based rough set approach using possibility and necessity measures, in: J.J. Alpigini, J.F. Peters, A. Skowron, N. Zhong (Eds.), Rough Sets and Current Trends in Computing, LNAI, vol. 2475, Springer-Verlag, Berlin, 2002, pp. 85–92.
[19] A. Hassanien, Fuzzy rough sets hybrid scheme for breast cancer detection, Image and Vision Computing 25 (2007) 172–183.
[20] E. Herrera-Viedma, F. Herrera, F. Chiclana, M. Luque, Some issues on consistency of fuzzy preference relations, European Journal of Operational Research 154 (2004) 98–109.
[21] T.P. Hong, T.T. Wang, S.L. Wang, et al, Learning a coverage set of maximally general fuzzy rules by rough sets, Expert Systems with Applications 19 (2000) 97–103.
[22] 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.
[23] 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.
[24] Q.H. Hu, D.R. Yu, Z.X. Xie, Neighborhood classifiers, Expert Systems with Applications 34 (2008) 866–876.
[25] R. Jensen, Q. Shen, Fuzzy-rough sets assisted attribute selection, IEEE Transactions on Fuzzy Systems 15 (2007) 73–89.
[26] R. Jensen, Q. Shen, Semantics-preserving dimensionality reduction: rough and fuzzy-rough -based approaches, IEEE Transactions on Knowledge and Data Engineering 16 (2004) 1457–1471.
[27] E.P. Klement, R. Mesiar, E. Pap, Triangular Norms, Kluwer Academic Publishers, 2001.
[28] J.W.T. Lee, D.S. Yeung, E.C.C. Tsang, Rough Sets and Ordinal Reducts, Soft Computing 10 (2006) 27–33.
[29] Y. Leung, M.M. Fischer, W.-Z. Wu, J.-S. Mi, A rough set approach for the discovery of classification rules in interval-valued information systems, International Journal of Approximate Reasoning 47 (2008) 233–246.
[30] T.-J. Li et al, Generalized fuzzy rough approximation operators based on fuzzy coverings, International Journal of Approximate Reasoning 48 (2008) 836–856.
[31] Y. Li, S.C.K. Shiu, S.K. Pal, Combining feature reduction and case selection in building CBR classifiers, IEEE Transactions on Knowledge and Data Engineering 18 (2006) 415–429.
[32] J.-S. Mi, W.-X. Zhang, An axiomatic characterization of a fuzzy generalization of rough sets, Information Sciences 160 (2004) 235–249.
[33] N.N. Morsi, M.M. Yakout, Axiomatics for fuzzy rough set, Fuzzy Sets Systems 100 (1998) 327–342.
[34] H.M. Moshkovich, A.I. Mechitov, D.L. Olson, Rule induction in data mining: effect of ordinal scales, Expert Systems with Applications 22 (2002) 303–311.
[35] Z. Pawlak, A. Skowron, Rough sets: some extensions, Information Sciences 177 (2007) 28–40.
[36] Z. Pawlak, A. Skowron, Rough sets and boolean reasoning, Information Sciences 177 (2007) 41–73.
[37] A. Petrosino, G. Salvi, Rough fuzzy set based scale space transforms and their use in image analysis, International Journal of Approximate Reasoning 41 (2006) 212–228.
[38] Y. Qian, J. Liang, C. Dang, Interval ordered information systems, Computers and Mathematics with Applications 56 (2008) 1994–2009.
[39] Y. Qian, J. Liang, D. Li, H. Zhang, C. Dang, Measures for evaluating the decision performance of a decision table in rough set theory, Information Sciences 178 (2008) 181–202.
[40] Y. Qian, J. Liang, C. Dang, Converse approximation and rule extracting from decision tables in rough set theory, Computers and Mathematics with Applications 55 (2008) 1754–1765.
[41] A.M. Radzikowska, E.E. Kerre, A comparative study of fuzzy rough sets, Fuzzy Sets and Systems 126 (2002) 137–155.
[42] R. Slowinski, S. Greco, B. Matarazzo, Axiomatization of utility, outranking and decision-rule preference models for multiple-criteria classification problems under partial inconsistency with the dominance principle, Control and Cybernetics 31 (2002) 1005–1035.
[43] V. Torra, Y. Narukawa, A view of averaging aggregation operators, IEEE Transactions on Fuzzy Systems 15 (2007) 1063–1067.
[44] Y.C. Tsai, C.H. Cheng, J.R. Chang, Entropy-based fuzzy rough classification approach for extracting classification rules, Expert Systems with Applications 31 (2) (2006) 436–443.
[45] S. Tsumoto, Mining diagnostic rules from clinical databases using rough sets and medical diagnostic model, Information Sciences 162 (2004) 65–80.
[46] Y.F. Wang, Mining stock price using fuzzy rough set system, Expert Systems with Applications 24 (1) (2003) 13–23.
[47] Y.-M. Wang, C. Parkan, Optimal aggregation of fuzzy preference relations with an application to broadband internet service selection, European Journal of Operational Research 187 (2008) 1476–1486.
[48] W.-Z. Wu, Attribute reduction based on evidence theory in incomplete decision systems, Information Sciences 178 (2008) 1355–1371.
[49] W.-Z. Wu, M. Zhang, H.-Z. Li, J.-S. Mi, Knowledge reduction in random information systems via Dempster–Shafer theory of evidence, Information Sciences 174 (2005) 143–164.
[50] W.-Z. Wu, J.-S. Mi, W.-X. Zhang, Generalized fuzzy rough sets, Information Sciences 151 (2003) 263–282.

[51] Z.S. Xu, Multiple-attribute group decision making with different formats of preference information on attributes, IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics 37 (2007) 1500–1511.
[52] Y.Y. Yao, Two views of the theory of rough sets in finite universes, International Journal of Approximate Reasoning 15 (1996) 291–317.
[53] Y. Sai, Y.Y. Yao, N. Zhong. Data analysis and mining in ordered information tables, in: Proceedings of the 2001 IEEE International Conference on Data Mining, pp. 497–504, 2001.
[54] Y. Yao, Y. Zhao, Attribute reduction in decision-theoretic rough set models, Information Sciences 178 (2008) 3356–3373.
[55] 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 Transactions on Fuzzy Systems 13 (2005) 343–361.
[56] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90 (1997) 111–127.
[57] L.A. Zadeh, Is there a need for fuzzy logic?, Information Sciences 178 (2008) 2751–2779