International Journal of Pure and Applied Mathematics
Volume 47 No. 1 2008, 89-100

A NEW SIMILARITY MEASURE ON FUZZY ROUGH SETS

Niu Qi[1], Zhang Chengyi[2 §]

[1]Department of Mathematics
Zhumadian Education College
Zhumadian, Henan, 463000, P.R. CHINA

[2]Department of Mathematics
Hainan Normal University
Haikou, Hainan, 571158, P.R. CHINA

Abstract: In this paper, we suggest some rules which are considered when we give a similarity measure for measuring of the degree of similarity between fuzzy rough sets and their elements. We propose a new similarity measures for measuring the degree of similarity between fuzzy rough sets and their elements. Finally, we illustrate the problem in the context of colorectal cancer diagnosis by similarity measure between fuzzy rough sets.

AMS Subject Classification: 03E72, 04A72
Key Words: fuzzy sets, rough sets, fuzzy rough sets, similarity measures

1. Introduction

The concept of fuzzy sets was introduced by Zadeh in his classical paper [13] in 1965, the concept of rough sets was proposed by Pawlak [12] in 1982. Although fuzzy sets and rough sets are methods to handle vague and inexact information, their starts and emphases are different. The fuzzy sets emphasize on the morbid definition of the boundary of sets, in which the relations of "belong to" and "not belong to" between elements and sets in the classical set theory are characterized by the membership degree; while rough sets use the approximation given the equivalent relation of the classical sets to research the indistinguishableness of elements. Two theories from different points of view may complement each

Received: May 3, 2008
§Correspondence author

other. Combining fuzzy sets with the rough sets, Nanda and Majumdar [10] proposed fuzzy rough sets (FRSs) in 1992.

Measures of similarity between fuzzy sets, as an important content in fuzzy mathematics, have gained attention from researchers for their wide applications in real world. Based on similarity measures that are very useful in some areas, such as pattern recognition, machine learning, decision making and market prediction, many measures of similarity between some fuzzy sets have been proposed and researched in recent years [8], [7]. For example, Chen [14], [4], [3] proposed some similarity functions to measure the degree of similarity between fuzzy sets and vague sets. Li and Cheng [2] also proposed similarity measures of intuitionistic fuzzy sets and applied similarity measures to pattern recognition. Zhang et al [9], [7] proposed similarity measures for measuring the degree of similarity between vague sets and fuzzy rough sets. In this paper, we first suggest some rules which is considered when we give a similarity measure between fuzzy rough sets and its elements. Moreover, we propose a new similarity measures between fuzzy rough sets and its elements. Finally, we illustrate the problem in the context of colorectal cancer diagnosis by similarity measure between fuzzy rough sets.

The rest of this paper is organized as follows. In Section 2, basic notions and definitions of fuzzy rough sets are reviewed. In Section 3, some existing measures of similarity are reviewed, some rules which are considered when we give a similarity measure between some fuzzy rough sets and its elements are proposed. In Section 4, we propose some new similarity measures between fuzzy rough sets. In Section 5, the proposed similarity measures are applied to deal with a problem related to pattern recognition. The final section is conclusions.

2. Fuzzy Rough Sets

Definition 2.1. (see [2]) Let U be a non-empty universe of discourse and R an equivalent relation on U, which is called an indistinguishable relation, $U/R = \{X_1, X_2, \cdots, X_n\}$ is all the equivalent class derived from R. $W = (U, R)$ are called an approximation space. $\forall X \subseteq U$, suppose $X_L = \{x \in U | [x] \subseteq X\}$ and $X_U = \{x \in U | [x] \cap X \neq \Phi\}$, a set pairs (X_L, X_U) are called a rough set in W, and denoted as $X = (X_L, X_U)$; X_L and X_U are the lower approximation and the upper approximation of X on W respectively.

Definition 2.2. (see [3]) Let S be the set of the whole rough sets, $X = (X_L, X_U) \in S$, then a fuzzy rough set $A = (A_L, A_U)$ in X can be described by a pair mapping $\mu_{A_L},\ \mu_{A_U}$

$$\mu_{A_L}:\ A_L \to [0,1], \quad \mu_{A_U}:\ A_U \to [0,1]\,.$$

Also $\mu_{A_L}(x) \le \mu_{A_U}(x)$. And then, a fuzzy rough set A in X could be denoted by

$$A = \{\langle x, \mu_{A_L}(x), \mu_{A_U}(x)\rangle | \forall x \in X\} = \{\langle \mu_{A_L}(x), \mu_{A_U}(x)\rangle | \forall x \in X\}\,. \quad (2.1)$$

The complement of A is denoted as

$$A^c = \{(x, \mu_{A_L^c}(x), \mu_{A_U^c}(x)) | \forall x \in X\} = \{(x,\ 1-\mu_{A_L}(x), 1-\mu_{A_U}(x)) | \forall x \in X\}\,. \quad (2.2)$$

The family of all fuzzy rough sets in X is denoted by $FR(X)$. The operations of fuzzy rough sets may be seen in [3].

Note 2.1. In [3], the complement of A is denoted as (6). Since $\mu_{A_L}(x) \le \mu_{A_U}(x),\ \forall x \in A_U$, then $1-\mu_{A_L}(x) \ge 1-\mu_{A_U}(x)$, it implies that

$$A^c = \{(x, 1-\mu_{A_L}(x), 1-\mu_{A_U}(x))\} \notin FR(X)\,.$$

This expresses that the definition in [3] cannot guarantee the closeness of $FR(X)$ for the complement operation. So we suggest that changing the following complement operation (2.3) to that (2.2).

$$A^c = \{(x, \mu_{A_L^c}(x), \mu_{A_U^c}(x)) | \forall x \in A_U\} = \{(x,\ 1-\mu_{A_L}(x), 1-\mu_{A_U}(x)) | \forall x \in A_U\}\,. \quad (2.3)$$

Note 2.2. The order relation in A is defined by the following condition:

$$x \le y \Leftrightarrow \mu_{A_L}(x) \le \mu_{A_L}(y)\ \textit{and}\ \mu_{A_U}(x) \le \mu_{A_U}(y)$$

3. The Similarity Measures and Some Rules

Many measures of similarity between fuzzy sets and intuitionistic fuzzy sets (IFSs) have been proposed and researched in recent years [5]-[12]. For example, in [8], Chen gave the similarity measure for measuring the degree of similarity between elements in a vague set (it is really an intuitionistic fuzzy set) as follows.

Definition 3.1. (see [8]) Let $x = [t_A(x), 1-f_A(x)]$ and $y = [t_A(y), 1-f_A(y)]$ be two fuzzy values in an IFS A. The degree of similarity between the fuzzy values x and y can be evaluated by the function M_C,

$$M_c(x,y) = 1 - \frac{1}{2}|S(x) - S(y)|\,, \quad (3.1)$$

where $S(x) = t_A(x) - f_A(x)$ and $S(y) = t_A(y) - f_A(y)$.

Definition 3.2. (see [**?**]) Let $x = [t_A(x), 1-f_A(x)]$ and $y = [t_A(y), 1-f_A(y)]$ be two fuzzy values in an intuitionistic fuzzy set A. The degree of similarity between the fuzzy values x and y can be evaluated by the function M_H

$$M_H(x, y) = 1 - \frac{1}{2}(|t_A(x) - t_A(y)| + |f_A(x) - f_A(y)|) \,. \tag{3.2}$$

In [7], Zhang gave the similarity measure between two fuzzy rough sets as follows.

Definition 3.3. (see [7]) Let A be a fuzzy rough set in X, $x = \langle \mu_{A_L}(x), \mu_{A_U}(x)\rangle$, $y = \langle \mu_{A_L}(y), \mu_{A_U}(y)\rangle$ the fuzzy rough values in A. The degree of similarity between the fuzzy rough values x and y can be evaluated by the function M_Z

$$M_Z(x, y) = 1 - \frac{1}{2}(|\mu_{A_L}(x) - \mu_{A_L}(y)| + |\mu_{A_U}(x) - \mu_{A_U}(y)|) \,. \tag{3.3}$$

Some Rules. Let $x = \langle \mu_{A_L}(x), \mu_{A_U}(x)\rangle$, $y = \langle \mu_{A_L}(y), \mu_{A_U}(y)\rangle$ and $z = \langle \mu_{A_L}(z), \mu_{A_U}(z)\rangle$ be the fuzzy rough values in a fuzzy rough set A. If M is the similarity measure for measuring the degree of similarity between elements in A, based on the background of fuzzy information handling, we give the definition of similarity measures between elements (or fuzzy rough values).

Definition 3.4. Let $A \in FR(X)$, $M : A \times A \to [0, 1]$, $(x, y) \to M(x, y)$. Then M is called the similarity measure on A and $M(x, y)$ is called the similarity degree between the fuzzy rough values $x = \langle \mu_{A_L}(x), \mu_{A_U}(x)\rangle$ and $y = \langle \mu_{A_L}(y), \mu_{A_U}(y)\rangle$, if M satisfies the following conditions:

Rule 1. $M(x, y) = M(y, x)$.

Rule 2. $M(x, y) = 0 \Leftrightarrow (x = [0, 0], y = [1, 1])$ or $(x = [1, 1], y = [0, 0])$; $M(x, y) = 1 \Leftrightarrow \mu_{A_L}(x), \mu_{A_L}(y)$ and $\mu_{A_U}(x), \mu_{A_U}(y)$.

Rule 3. If $\mu_{A_L}(x) \le \mu_{A_L}(y) \le \mu_{A_L}(z)$ and $\mu_{A_U}(x) \le \mu_{A_U}(y) \le \mu_{A_U}(z)$, then

$$M(x, z) \le \min\{M(x, y), M(y, z)\} \,. \tag{3.4}$$

Rule 4. $M(x, y) = M(x^c, y^c)$.

Rule 5. $\forall x \in X,\ M(x, y) = M(x, z) \Rightarrow M(y, z) = 1$.

4. The Similarity Measure between Fuzzy Rough Sets and its Elements

Definition 4.1. Let $A \in FR(X)$ and $x = \langle \mu_{A_L}(x), \mu_{A_U}(x)\rangle \in A$. Then:

(1) $T_x = \mu_{A_U}(x) - \mu_{A_L}(x)$ is called the degree of indeterminacy of the element $x \in A$.

(2) $\delta_x = \mu_{A_L}(x) + T_x\mu_{A_L}(x)$ is called the degree of favor $x \in A$.

(3) $\alpha_x = 1 - \mu_{A_U}(x) + T_x(1 - \mu_{A_U}(x))$ is called the degree of against $x \in A$.

Notes. (1) The larger the value of T_x, the more the degree of unknown for x. Especially, if $T_x = 1$, we know nothing for x; if $\forall x \in A,\ T_x = 0$, then the fuzzy rough set A is a fuzzy set; if $\forall x \in A,\ \mu_{A_U}(x) = \mu_{A_L}(x) = 1(0)$, then the fuzzy rough set A is a common set.

(2) The prior knowledge is considered when defining δ_x and α_x, it can be interpreted by the voting model. A fuzzy rough value $[0.4, 0.8]$ can be interpreted as "the vote for resolution is 4 in favor, 2 against, and 4 abstention". Then $\delta_x = 0.4 + 0.4(1 - 0.4 - 0.2) = 0.56$ can be interpreted as "considering the vote for resolution as above, besides 4 in favor, it is possible that there is 0.4×4 favor in the 4 abstention". Similarly, $\alpha_x = 0.2 + 0.2(1 - 0.4 - 0.2) = 0.28$ can be interpreted as "considering the vote for resolution as above, besides 2 against, it is possible that there is 0.2×4 against in the 4 abstention".

Definition 4.2. Let $A \in FRS(X)$ and T_x, δ_x, α_x, as above. $x = \langle \mu_{A_L}(x), \mu_{A_U}(x) \rangle$ and $y = \langle \mu_{A_L}(y), \mu_{A_U}(y) \rangle$ in A. The similarity degree between x and y can be evaluated by the function M,

$$M(x, y) = 1 - \frac{1}{2}(\delta_{xy} + \alpha_{xy})\,, \tag{4.1}$$

where $\delta_{xy} = |\delta_x - \delta_y|$ and $\alpha_{xy} = |\alpha_x - \alpha_y|$.

We will prove this similarity degree satisfying the Rules 1-5 in Definition 3.4.

Proposision 4.1. *Let $A \in FRS(X)$ and M be as above. $x = \langle \mu_{A_L}(x), \mu_{A_U}(x) \rangle$, $y = \langle \mu_{A_L}(y), \mu_{A_U}(y) \rangle$ and $z = \langle \mu_{A_L}(z), \mu_{A_U}(z) \rangle \in A$. If $x \le y \le z$, then $M(x, z) \le \min\{M(x, y), M(y, z)\}$.*

Proof. Let $\mu_{A_L}(x) = \mu_{A_L}(y) - u$ and $\mu_{A_U}(x) = \mu_{A_U}(y) - v$, where $u > 0$ and $v > 0$, Then $T_x = T_y + (v - u)$.

$\delta_y - \delta_x = u(1 + T_x) + \mu_{A_L}(y)(v - u) \ge u\mu_{A_L}(y) + \mu_{A_L}(y)(v - u) = \mu_{A_L}(y)v \ge 0\,,$
so $\delta_z - \delta_x = \delta_z - \delta_y + \delta_y - \delta_z \ge \max\{\delta_z - \delta_y, \delta_y - \delta_x\}$, by the nonnegativity of $\delta_z - \delta_y$ and $\delta_y - \delta_x$, we have $\delta_{xz} \ge \max\{\delta_{xy}, \delta_{yz}\}$, then $M(x, z) \le \min\{M(x, y), M(y, z)\}$. □

Proposision 4.2. *Let $A \in FRS(X)$ and M be as above. Then:*

(1) $M(x, y) = 0 \Leftrightarrow (x = \langle 0, 0 \rangle$ and $y = \langle 1, 1 \rangle)$ or $(x = \langle 1, 1 \rangle$ and $y = \langle 0, 0 \rangle)$.

(2) $M(x, y) = 1 \Leftrightarrow \mu_{A_L}(x) = \mu_{A_L}(y)$ and $\mu_{A_U}(x) = \mu_{A_U}(y)$.

Proof. (1) Firstly, since $0 \le \delta_x = \mu_{A_L}(x) + (\mu_{A_U}(x) - \mu_{A_L}(x))\mu_{A_L}(x) \le \mu_{A_L}(x) + (\mu_{A_U}(x) - \mu_{A_L}(x)) = \mu_{A_U}(x) \le 1$, then $\delta_{xy} = |\delta_x - \delta_y| \le 1$. Similarly, we have $\alpha_{xy} \le 1$.

Secondly, $M(x, y) = 0 \Leftrightarrow \delta_{xy} + \alpha_{xy} = 2 \Leftrightarrow \delta_{xy} = \alpha_{xy} = 1$.

If $\alpha_{xy} = |\alpha_x - \alpha_y| = 1$, then from $0 \le \alpha_x \le 1$ and $0 \le \alpha_y \le 1$ implies $\alpha_x = 1$ and $\alpha_y = 0$ or $\alpha_x = 0$ and $\alpha_y = 1$. Similarly, if $\delta_{xy} = |\delta_x - \delta_y| = 1$, we get $\delta_x = 1$ and $\delta_y = 0$ or $\delta_x = 0$ and $\delta_y = 1$.

When $\alpha_x = 1$, because $\alpha_x = 1 - \mu_{A_U}(x) + T_x(1 - \mu_{A_U}(x)) \le 1 - \mu_{A_U}(x) + T_x \le 1 - \mu_{A_L}(x)$, thus $\mu_{A_L}(x) = 0$; Since $0 \le \mu_{A_U}(x) \le 1$, $1 = \alpha_x = (1 - \mu_{A_U}(x))(1 + \mu_{A_U}(x)) = 1 - (\mu_{A_U}(x))^2$, it implies that $\mu_{A_U}(x) = 0$. When $\alpha_y = 0$, we have $\mu_{A_U}(y) = 1$ by nonnegativity of $\mu_{A_U}(y)$ and T_y. So, we have $\mu_{A_U}(x) = \mu_{A_L}(x) = \delta_x = 0$ and $\mu_{A_U}(y) = 1$ when $\alpha_x = 1$ and $\alpha_y = 0$.

Similarly, we have $\mu_{A_U}(y) = \mu_{A_L}(y) = \delta_y = 0$ and $\mu_{A_U}(x) = 1$ when $\alpha_x = 0$ and $\alpha_y = 1$.

By same discussion, when $\delta_x = 0$ and $\delta_y = 1$, we have $\mu_{A_L}(y) = 0$; when $\delta_x = 1$ and $\delta_y = 0$, we have $\mu_{A_L}(x) = 1$.

Thus, $\delta_{xy} = \alpha_{xy} = 1 \Leftrightarrow (x = \langle 0, 0\rangle$ and $y = \langle 1, 1\rangle)$ or $(x = \langle 1, 1\rangle$ and $y = \langle 0, 0\rangle)$.

(2) Fist of all, $M(x, y) = 1 \Leftrightarrow \delta_{xy} = \alpha_{xy} = 0 \Leftrightarrow \delta_x = \delta_y$ and $\alpha_x = \alpha_y \Leftrightarrow (\mu_{A_L}(x) + 1 - \mu_{A_U}(x))(1 + T_x) = (\mu_{A_L}(y) + 1 - \mu_{A_U}(y))(1 + T_y)$ and $(\mu_{A_L}(x) - 1 + \mu_{A_U}(x))(1 + T_x) = (\mu_{A_L}(y) - 1 + \mu_{A_U}(y))(1 + T_y) \Rightarrow (1 - T_x)(1 + T_x) = (1 - T_y)(1 + T_y) \Rightarrow T_x = T_y \Rightarrow \mu_{A_U}(x) - \mu_{A_L}(x) = \mu_{A_U}(y) - \mu_{A_L}(y)$

Second, from $T_x = T_y$ and $(\mu_{A_L}(x) - 1 + \mu_{A_U}(x))(1 + T_x) = (\mu_{A_L}(y) - 1 + \mu_{A_U}(y))(1 + T_y)$ implies $\mu_{A_U}(x) + \mu_{A_L}(x) = \mu_{A_U}(y) + \mu_{A_L}(y)$, thus we have $\mu_{A_L}(x) = \mu_{A_L}(y)$ and $\mu_{A_U}(x) = \mu_{A_U}(y)$. It is clear that $\delta_{xy} = \alpha_{xy} = 0$ if $\mu_{A_L}(x) = \mu_{A_L}(y)$ and $\mu_{A_U}(x) = \mu_{A_U}(y)$. □

Proposision 4.3. *Let $A \in FRS(X)$ and M be as above. Then $M(x, y) = M(y, x)$ and $M(x, y) = M(x^c, y^c)$.*

Proposision 4.4. *Let $A \in FRS(X)$ and M be as above. If $M(x, y) = M(x, z)$ for all $x \in X$, then $M(y, z) = 1$.*

Proof. If for all $x \in X$, $M(x, y) = M(x, z)$, then $|\delta_x - \delta_y| + |\alpha_x - \alpha_y| = |\delta_x - \delta_z| + |\alpha_x - \alpha_z|$. Let $x \in X$ and x satisfies $\alpha_x = 1$, by the proof of above proposision, we have $\delta_x = 0$, thus $\delta_y + |1 - \alpha_y| = \delta_z + |1 - \alpha_z|$. Then $\delta_y - \alpha_y = \delta_z - \alpha_z$. Let $x \in X$ such that $t_x = f_x = 0$, then $\delta_x = \alpha_x = 0$, it implies that $\delta_y + \alpha_y = \delta_z + \alpha_z$. Thus $\delta_y = \delta_z$ and $\alpha_y = \alpha_z$, so $M(y, z) = 1$. □

Definition 4.3. Let $A, B \in FRS(X)$, $X = \{x_1, x_2, \cdots, x_n\}$. If $FRS_A(x) =$

$\langle \mu_{A_L}(x), \mu_{A_U}(x) \rangle$ is the fuzzy rough value of x in A and

$$FRS_B(x) = \langle \mu_{B_L}(x), \mu_{B_U}(x) \rangle$$

is the fuzzy rough value of x in B. Then the degree of similarity between the vague sets A and B can be evaluated by the function S.

$$\begin{aligned} S(A, B) &= \frac{1}{n} \sum_{i=1}^{n} M(FRS_A(x_i), FRS_B(x_i)) \\ &= \frac{1}{n} \sum_{i=1}^{n} (1 - \frac{\delta_A(x_i) - \delta_B(x_i)}{2} - \frac{\alpha_A(x_i) - \alpha_B(x_i)}{2}) \\ &= 1 - \frac{1}{2n} \sum_{i=1}^{n} (|\delta_A(x_i) - \delta_B(x_i)| + |\alpha_A(x_i) - \alpha_B(x_i)|), \quad (4.2) \end{aligned}$$

where $\delta_A(x_i) = \mu_{A_L}(x_i) + (\mu_{A_U}(x_i) - \mu_{A_L}(x_i))\mu_{A_L}(x_i)$ and

$$\alpha_A(x_i) = 1 - \mu_{A_U}(x_i) + (\mu_{A_U}(x_i) - \mu_{A_L}(x_i))(1 - \mu_{A_U}(x_i)).$$

It is obvious that $S(A, B) \in [0, 1]$. The larger the value of $S(A, B)$, the more the similarity between A and B.

The following conclusions are obvious.

Proposision 4.5. $S(A, B) = S(B, A),\ S(A, B) = S(A^c, B^c)$.

Proposision 4.6. $S(A, B) = 0 \Leftrightarrow$ *(*$A = \sum_{i=1}^{n} < 0, 0 > /x_i$ *and* $B = \sum_{i=1}^{n} < 1, 1 > /x_i$*) or (*$A = \sum_{i=1}^{n} < 1, 1 > /x_i$ *and* $B = \sum_{i=1}^{n} < 0, 0 > /x_i$*).*

Proposition 4.7. $S(A, B) = 1 \Leftrightarrow \mu_{A_L}(x_i) = \mu_{B_L}(x_i)$ *and* $\mu_{A_U}(x_i) = \mu_{B_U}(x_i), \forall x_i \in X$.

Proposition 4.8. $\forall A, B, C \in V(X),\ A \subseteq B \subseteq C \Rightarrow S(A, C) \leq S(A, B) \wedge S(B, C)$.

Thus, S is a strong measure of similarity on $FRS(X)$.

Similarity, we can define the weighting similarity measures between IFSs.

Let $A, B \in FRS(X),\ X = \{x_1, x_2, x_3, \cdots, x_n\}$, where

$$A = \sum_{i=1}^{n} [\mu_{A_L}(x_i), \mu_{A_U}(xi)]/x_i, \quad B = \sum_{i=1}^{n} [\mu_{B_L}(x_i), \mu_{B_U}(xi)]/x_i, \qquad x_i \in X.$$

Assume that the weight of the element x_i is w_i, where $w_i \in [0, 1]$ and $1 \leq i \leq n$, then the degree of similarity between the vague sets A and B can be evaluated by the weighting function W

$$W(A, B) = \sum_{i=1}^{n} w_i M(FRS_A(x_i), FRS_B(x_i)) / \sum_{i=1}^{n} w_i$$

$$= \sum_{i=1}^{n} w_i (1 - \frac{|\delta_A(x_i) - \delta_B(x_i)|}{2} - \frac{|\alpha_A(x_i) - \alpha_B(x_i)|}{2}) / \sum_{i=1}^{n} w_i \,, \quad (4.3)$$

where $W(A,B) \in [0,1]$. The largest value of $W(A,B)$, implies more similarity between A and B.

5. Applications

The measures of similarity between the FRSs, can be used to measure the importance of a feature in a given classification task. Here we illustrate this problem in the context of colorectal cancer diagnosis. In this example, we study the association between the key prognostic factors and the outcomes of the patients who are undergoing the follow-up program of the colorectal cancer.

The colorectal cancer occurs frequently in the developed countries. The colorectal cancer forms initially in the mucosa lining of bowel. In most cases, the first step in the formation of a colorectal cancer is the appearance of polyps. When the abnormal cells within the polyps begin to spread and invade through normal tissue, polyps become cancer growths. If no proper treatment is adopted, then the cancer can spread beyond the skin and the underlying tissues of the bowel wall, and eventually the cancer may spread to the distant sites like liver. To express the condition of the patient, the following for possible outcomes are used: Well, recurrence, metastasis and both (i.e., recurrence and metastasis simultaneously). The main treatment for the colorectal cancer is the surgical removal of the tumor, while the survival of a patient with the colorectal cancer is dependent on four fundamental factors: (a) the biology of that individual's malignancy, (b) the immune response to the tumor, (c) the time in the cancer patient's life history when the diagnosis is made, and (d) the adequacy of the treatment. About 50 patients eventually die from the local recurrence and/or distant metastasis within 5 years after the curative resection. Therefore, it is important to detect or predict the recurrent or metastasis tumor in the follow-up so that the appropriate therapy is prescribed to increase the chance of survival.

The patient, who is in the follow-up program, may fall into any of the following states: Metastasis, recurrence, bad and well. If the state of a particular patient can be correctly decided, then the state information can be utilized to choose an appropriate treatment. A physician can subjectively judge the belongingness of each patient in the output classes.

Let A be an attribute set of a patient, for the convenient of discussion, we

	a	b	c	d	e
A_1	[0.4,0.6]	[0.3,0.7]	[0.5,0.9]	[0.5,0.8]	[0.6,0.8]
A_2	[0.2,0.4]	[0.3,0.5]	[0.2,0.7]	[0.7,0.9]	[0.8,1]
A_3	[0.1,0.1]	[0,0]	[0.2,0.3]	[0.1,0.2]	[0.2,0.2]
A_4	[0.8,0.8]	[0.9,1]	[1,1]	[0.7,0.8]	[0.6,0.6]

Table 1: Attribute sets of the sample

adopt the main 5 attributes and quantify the attribute, respectively denoted a, b, c, d, e; as these attributes usually are language variable, for every attribute, we establish FRSs function by fuzzy method or probability method, and obtain their attribute values. Let mode $A = \{[0.3, 0.5], [0.4, 0.6], [0.6, 0.8], [0.5, 0.9], [0.9, 1]\}$, A_1, A_2, A_3 and A_4 are the attribute sets of the samples denoted metastasis, recurrence, bad and well, shown in Table 1.

Using the method in Definition 3.3, we compute the similarity measure between the patient A and the sample A_1, A_2, A_3 and A_4: $A = \{[0.3, 0.5], [0.4, 0.6], [0.6, 0.8], [0.5, 0.9], [0.9, 1]\}$

$$S(A, A_1) = 1 - \frac{1}{10}(|0.3 - 0.4| + |0.5 - 0.6| + |0.4 - 0.3| + |0.6 - 0.7|$$
$$+|0.6-0.5|+|0.8-0.9|+|0.5-0.5|+|0.9-0.8|+|0.9-0.6|+|1-0.8|) = 0.88$$

$$S(A, A_2) = 1 - \frac{1}{10}(|0.3 - 0.2| + |0.5 - 0.4| + |0.4 - 0.3| + |0.6 - 0.5|$$
$$+|0.6-0.2|+|0.8-0.7|+|0.5-0.7|+|0.9-0.9|+|0.9-0.8|+|1-1|) = 0.88$$

$$S(A, A_3) = 1 - \frac{1}{10}(|0.3 - 0.1| + |0.5 - 0.1| + |0.4 - 0| + |0.6 - 0|$$
$$+|0.6-0.2|+|0.8-0.3|+|0.5-0.1|+|0.9-0.2|+|0.9-0.2|+|1-0.2|) = 0.49$$

$$S(A, A_4) = 1-\frac{1}{10}(|0.3-0.8|+|0.5-0.8|+|0.4-0.9|+|0.6-1|+|0.6-1|+|0.8-1|$$
$$+ |0.5 - 0.7| + |0.9 - 0.8| + |0.9 - 0.6| + |1 - 0.6|) = 0.67\,.$$

Since $S(A, A_1) = S(A, A_2)$, we could not tell A is a metastasis patient or a recurrence patient. It shows that the method in Definition 3.3 could not distinguish A is similar to A_1 or A_2. While using the method in Definition 4.2, A could be distinguished correctly, so he may be treated properly. The process is followed:

	a	b	c	d	e
$\mu_{A_{1L}}$	0.4	0.3	0.5	0.5	0.6
$\mu_{A_{1U}}$	0.6	0.7	0.9	0.8	0.8
T_{A_1}	0.2	0.4	0.4	0.3	0.2
δ_{A_1}	0.48	0.42	0.70	0.65	0.72
α_{A_1}	0.48	0.42	0.14	0.26	0.24

	a	b	c	d	e
$\mu_{A_{2L}}$	0.2	0.3	0.2	0.7	0.8
$\mu_{A_{2U}}$	0.4	0.5	0.7	0.9	1
T_{A_2}	0.2	0.2	0.5	0.2	0.2
δ_{A_2}	0.24	0.36	0.30	0.84	0.96
α_{A_2}	0.72	0.60	0.45	0.12	0.0

	a	b	c	d	e
$\mu_{A_{3L}}$	0.1	0.0	0.2	0.1	0.2
$\mu_{A_{3U}}$	0.1	0.0	0.3	0.2	0.2
T_{A_3}	0.0	0.0	0.1	0.1	0.0
δ_{A_3}	0.1	0.0	0.22	0.11	0.20
α_{A_3}	0.3	1.0	0.77	0.88	0.80

	a	b	c	d	e
$\mu_{A_{4L}}$	0.8	0.9	1	0.7	0.6
$\mu_{A_{4U}}$	0.8	1	1	0.8	0.6
T_{A_4}	0.0	0.1	0.0	0.1	0.0
δ_{A_4}	0.8	0.99	1.0	0.77	0.60
α_{A_4}	0.2	0.0	0.0	0.22	0.40

	a	b	c	d	e
μ_{A_L}	0.3	0.4	0.6	0.5	0.9
μ_{A_U}	0.5	0.6	0.8	0.9	1.0
T_A	0.2	0.2	0.2	0.4	0.1
δ_A	0.36	0.48	0.72	0.70	0.99
α_A	0.60	0.48	0.24	0.14	0.0

$$S'(A, A_1) = 1 - \frac{1}{10}(|0.36 - 0.48| + |0.6 - 0.48| + |0.48 - 0.42| + |0.48 - 0.42| + |0.72 - 0.7| + |0.24 - 0.14| + |0.7 - 0.65| + |0.14 - 0.26| + |0.99 - 0.72| + |0 - 0.24|) = 0.884$$

$$S'(A, A_2) = 1 - \frac{1}{10}(|0.36 - 0.24| + |0.6 - 0.72| + |0.48 - 0.36| + |0.48 - 0.60| + |0.72 - 0.3| + |0.24 - 0.45| + |0.7 - 0.84| + |0.14 - 0.12| + |0.99 - 0.96| + |0 - 0|) = 0.87$$

$$S'(A, A_3) = 1 - \frac{1}{10}(|0.36 - 0.1| + |0.6 - 0.3| + |0.48 - 0| + |0.48 - 0.10| + |0.72 - 0.22| + |0.24 - 0.77| + |0.7 - 0.11| + |0.14 - 0.88| + |0.99 - 0.2| + |0 - 0.8|) = 0.449$$

$$S'(A, A_4) = 1 - \frac{1}{10}(|0.36 - 0.8| + |0.6 - 0.2| + |0.48 - 0.99| + |0.48 - 0.0| + |0.72 - 1.0| + |0.24 - 0.0| + |0.7 - 0.77| + |0.14 - 0.22| + |0.99 - 0.6| + |0 - 0.4|) = 0.671\,.$$

Basing the above result, mode A is supposed to a recurrence patient.

6. Conclusions

The fuzzy rough set is one kind of concepts representing inexact knowledge. Measures of similarity between fuzzy rough sets, as an important content in fuzzy mathematics, have gained attention from researchers for their wide applications in real world. Based on similarity measures that are very useful in some areas, such as pattern recognition, machine learning, decision making and market prediction, many measures of similarity between these sets have been proposed and researched in recent years. Constantly looking for the better similarity measure method is a pursuing of fuzzy set and rough set theory. However, none of any a similarity measure methods are all-powerful, and all have the localization of its usage. We think that should according to the applied background that the indetermination information handle, bring up some basic rule when we define some similarity measures. We proposed a new similarity measure between fuzzy rough sets and its elements. Moreover, we give some new similarity measures between these fuzzy sets. Because of the proposed similarity measures have some good properties, it can provide a useful way for measuring the similarity between fuzzy rough sets. At the same time, It will be a very meaningful research topic that investigating the invariance properties of pattern recognition under the different similarity measure methods.

Acknowledgements

The authors are very grateful to the referees for their useful comments.

Project supported by NSF of P.R. China (60364001, 70461001) and Hainan (80401).

References

[1] M. Banerjee, S.K. Pal, Roughness of a fuzzy set, *Information Science*, **93** (1996), 235-246.

[2] Shyi-Ming Chen, J.M. Tan, Handling multi-criteria fuzzy decision-making problems based on vague sets, *Fuzzy Sets Systems*, **67** (1994), 163-172.

[3] Shyi-Ming Chen et al, A comparison of similarity measures of fuzzy values, *Fuzzy Sets Systems*, **72** (1995), 79-89.

[4] Shyi-Ming Chen, Measures of similarity between vague sets, *Fuzzy Sets and Systems*, **74** (1995), 217-223.

[5] Zhang Chengyi et al, On measures of similarity between fuzzy rough sets, *International Journal of Pure and Applied Mathematics*, **10**, No. 4, 451-460.

[6] Zhang Chengyi et al, The measurees of similarity between vague sets, *Computer Engineering and Applications*, **39**, No 17 (2003), 92-94, In Chinese.

[7] Dug Hun Hong, Chul Kim, A note on similarity measures between vague sets and between elements, *Information Sciences*, **115** (1999), 83-96.

[8] L.K. Hyung et al, Similarity measure between fuzzy sets and between elements, *Fuzzy Sets and Systems*, **62** (1994), 291-293.

[9] D. Li, C. Cheng, New similarity measures of intuitionistic fuzzy sets and application to pattern recognition, *Pattern Recognition Lett.*, **23**, No-s: 1-3, 221-225.

[10] S. Nana, S. Majumdar, Fuzzy rough sets, *Fuzzy Sets and Systems*, **45** (1992), 157-160.

[11] C.P. Pappis, N.I. Karacapilidis, A comparative assessment of measures of similarity of fuzzy values, *Fuzzy Sets and Systems*, **56** (1993), 171-174.

[12] Z. Pawlak, Rough Sets, *Inter. J. Inform. Comput. Sci.*, **11**, No. 5 (1982), 341-356.

[13] L.A. Zadeh, Fuzzy sets, *Inform. and Control*, **8** (1965), 338-356.

[14] R. Zwick et al, Measures of similarity among fuzzy sets: a comparative analysis, *Internat. J. Approx. Reason.*, **1** (1987), 221-242.