Appl. and Comput. Math., V.8, N.2, 2009, pp.203-215

CLASSIFICATION BASED ON SIMILARITY AND DISSIMILARITY THROUGH EQUIVALENCE CLASSES

M.KESHAVARZI †, M.A.DEHGHAN ‡, M.MASHINCHI ††, §

Abstract. In this paper we first briefly review the literature on similarity relations, and then introduce a dissimilarity relation between a pair of elements in a specific domain. Since in some cases recognizing dissimilarity is easier than similarity, we try to find a connection between these relations based on specific functions.

Keywords: fuzzy sets, similarity relations, dissimilarity relations.

AMS Subject Classification: 62H30, 03E20, 03B52.

1. Introduction

The most primary observation we can make when studying a group of objects or phenomena is that some are similar and others dissimilar. Similar objects can be grouped together to form a class, and we define a class as a set of similar objects. We always compare two objects with respect to a frame of reference, such as basic characteristics, context, or point of view. In other words, background information, or the existence of other classes, will affect the way objects are compared. The concept of similarity has been studied by many researchers. Similarity-based clustering is described in [5, 12, 26]. Chakraborty and Das [6], Valverede, Trillas and Jakas [11, 25, 27], Ovchinikov [16, 17] have widely studied the similarities in various contexts.

The notion of similarity originated in psychology and was established to determine why and how entities are grouped to categories, and why some categories are comparable to each other while others are not [9, 10]. The main challenge in semantic similarity measurement is the comparison of meanings. Human judgments of similarity have been subject to research in psychology for more than fifty years [9]. Different approaches to modelling similarity including feature-based, network-based, and geometric approaches have been developed. More recently, the Artificial Intelligence (AI) community started investigations on computational similarity models as a new method for information retrieval [23]. The Matching Distance Similarity Measure (MDSM) [24] was the first similarity-based model that has been developed specifically for the geospatial domain. V. Loia *et al.* in [14] used similarity relations in an internet e-mail application which is concerned with finding people who are interested in receiving a particular message via e-mail. Also similarity measures have been used extensively in text summarization as in [1] where the performance of different similarity measures have been evaluated in the context of document summarization. In this paper we basically study similarity and dissimilarity relations based on equivalence classes, but it is time worthy to pay attention to the work done on the application of similarity measures to fuzzy sets[28], which is an important tool in fuzzy

†Faculty of Mathematics, Vali-e-Asr University, Rafsanjan, Iran, e-mail: mkeshavarzi@mail.vru.ac.ir
‡Faculty of Mathematics, Vali-e-Asr University, Rafsanjan, Iran, e-mail: dehghan@mail.vru.ac.ir
††Faculty of Mathematics and Computer Science, Shahid Bahonar University, Kerman, Iran, e-mail: mashinchi@mail.uk.ac.ir
§*Manuscript received 3 July, 2008.*

mathematics, fuzzy economics [2], decision making, market prediction, and pattern recognition [7, 18, 29]. The measure of similarity of fuzzy sets has been proposed by Zwick *et al.* in [29]. Rezaei *et al.* in [21, 22] proposed a new similarity measure between fuzzy sets and extended it to define two other similarity measures. Similarity/dissimilarity topics have also been extensively used in the study of Intuitionistic Fuzzy Sets (IFS)[8, 15], introduced by Atanassov [3, 4] in 1983. IFS form an extension of fuzzy sets. Rezaie and Mukaidono in [19, 20, 22] proposed new similarity measures for IFS. They also studied the relationship between similarity and dissimilarity measures of two IFS sets.

Similarity and dissimilarity are closely related. Two objects that have large similarity have small dissimilarity, both of which imply a close resemblance. Finding functions that relate these two concepts is our primary goal. The construction of the paper is as follows. After the introduction, in Section 2, the theoretical aspects of similarity relations are discussed. In Section 3, we study dissimilarity relations and we prove related propositions. In Section 4, we show the relationship between similarity and dissimilarity and in Section 5, we propose a function that relates similarities and dissimilarities. Section 6 concludes the paper.

2. Similarity relation

A similarity relation is a mathematical notion that provides a way to manage alternative instances of an entity that can be considered "equal" to a given degree [13, 28]. The following definitions and proposition are due to V. Loia *et al.* [14].

Definition 2.1. *A similarity on a domain U is a fuzzy subset $S : U \times U \longrightarrow [0, 1]$ such that the following properties hold:*

(i) $S(x, x) = 1$ *for any* $x \in U$ *(reflexivity)*
(ii) $S(x, y) = S(y, x)$ *for any* $x, y \in U$ *(symmetry)*
(iii) $S(x, z) \geq S(x, y) \wedge S(y, z)$ *for any* $x, y, z \in U$*, where* $\wedge$ *is a minimum operator (transitivity).*
We say that S is strict if the following implication is also true.
(iv) $S(x, z) = 1 \Rightarrow x = z$.

Definition 2.2. *Let $S : U \times U \longrightarrow [0, 1]$ be a similarity on U. Then, for any $\lambda \in [0, 1]$, the relation $\cong_{S,\lambda}$ in U defined by*

$$x \cong_{S,\lambda} y \Leftrightarrow S(x, y) \geq \lambda$$

is called a $\lambda - \text{cut}$ of S.
The notion of $\lambda - \text{cuts}$ allows us to define a similarity by means of suitable family of equivalence relations according to the following result.

Proposition 2.1. *a) Let S be a similarity in a domain U and, for any $\lambda \in [0, 1]$ let $\cong_{S,\lambda}$ be the $\lambda - \text{cut}$ of S. Then,$\{\cong_{S,\lambda}\}_{\lambda \in [0,1]}$ is a family of equivalence relations such that:*

(i) *for any* μ *and* λ *in* $[0, 1], \lambda \leq \mu \Rightarrow \cong_{S,\lambda} \supseteq \cong_{S,\mu}$
(ii) *for any* μ *in* $[0, 1], \bigcap_{\lambda \leq \mu} \cong_{S,\lambda} = \cong_{S,\mu}$.

b) Let $\{\cong_\lambda\}_{\lambda \in [0,1]}$ be a family of equivalence relations satisfying conditions (i) and (ii) above, then the relation S defined by setting

$$S(x, y) = Sup\{\lambda \in [0, 1] | x \cong_\lambda y\}$$

is a similarity whose family of $\lambda - \text{cuts}$ is equal to the family $\{\cong_\lambda\}_{\lambda \in [0,1]}$.

3. Dissimilarity relation

Dissimilarity relation is a mathematical notion that provides a way to manage alternative instances of an entity that can be considered "different" to a given degree.

Definition 3.1. *A dissimilarity on a domain U is a fuzzy subset $D : U \times U \longrightarrow [0, 1]$ such that the following properties hold:*

(i) $D(x, x) = 0$ *for any* $x \in U$ *(reflexivity)*
(ii) $D(x, y) = D(y, x)$ *for any* $x, y \in U$ *(symmetry)*
(iii) $D(x, z) \leq D(x, y) \vee D(y, z)$ *for any* $x, y, z \in U$,
where $\vee$ *is a maximum operator (trasitivity). We say that* D *is strict if the following implication also holds.*
(iv) $D(x, z) = 0 \Rightarrow x = z$.

Definition 3.2. *Let $D : U \times U \longrightarrow [0, 1]$ be a dissimilarity on U. Then, for any $\lambda \in [0, 1]$, the relation $\cong_{D,\lambda}$ in U defined by*

$$x \cong_{D,\lambda} y \Leftrightarrow D(x, y) \leq \lambda \tag{1}$$

is called a $\lambda - \mathrm{cut}$ *of* D.

Proposition 3.1. *Let U be a domain and $D : U \times U \longrightarrow [0, 1]$ a fuzzy relation in U. Then, for any $\lambda \in [0, 1]$, the relation $\cong_{D,\lambda}$ in U defined by*

$$x \cong_{D,\lambda} y \Leftrightarrow D(x, y) \leq \lambda$$

is a dissimilarity relation which can be defined by means of a suitable family of equivalence relations.

Proof. We investigate the reflexivity, symmetric, and transitivity properties of D.

- Reflexivity: According to the Definitions 3.1 and 3.2, since $D(x, x) = 0$, it is obvious that $D(x, x) \leq \lambda$ for any λ. That is $x \cong_{D,\lambda} x$.
- Symmetry: To prove $x \cong_{D,\lambda} y$ implies that $y \cong_{D,\lambda} x$. Note that:
$x \cong_{D,\lambda} y \Rightarrow D(x, y) \leq \lambda \Rightarrow D(y, x) \leq \lambda \Rightarrow x \cong_{D,\lambda} y$.
- Transitivity: To prove $x \cong_{D,\lambda} y$ and $y \cong_{D,\lambda} z$ imply $x \cong_{D,\lambda} z$. Note that:
$x \cong_{D,\lambda} y$ implies $D(x, y) \leq \lambda$ and $y \cong_{D,\lambda} z$ implies $D(y, z) \leq \lambda$. According to (iii) of Definition 3.1, since $D(x, y) \leq \lambda$ and $D(y, z) \leq \lambda$. Consequently $D(x, z) \leq \lambda$ and hence $x \cong_{D,\lambda} z$.

Proposition 3.2. *a) Let D be a dissimilarity in a domain U and, for any $\lambda \in [0, 1]$, let $\cong_{D,\lambda}$ be the $\lambda - \mathrm{cut}$ of D. Then,$\{\cong_{D,\lambda}\}_{\lambda \in [0,1]}$ is a family of equivalence relations such that:*

(i) *for any* μ *and* λ *in* $[0, 1], \lambda \leq \mu \Rightarrow \cong_{D,\lambda} \subseteq \cong_{D,\mu}$.
(ii) *for any* μ *in* $[0, 1], \bigcap_{\lambda \geq \mu} \cong_{D,\lambda} = \cong_{D,\mu}$.

b) Let $\{\cong_\lambda\}_{\lambda \in [0,1]}$ be a family of equivalence relations satisfying conditions (i) and (ii)above, then the relation D defined by

$$D(x, y) = Inf\{\lambda \in [0, 1] | x \cong_\lambda y\}$$

is a dissimilarity whose family of $\lambda - \mathrm{cuts}$ *is* $\{\cong_\lambda\}_{\lambda \in [0,1]}$.

Proof. (i) Let $\lambda \leq \mu$ and let $(x, y) \in \cong_{D,\lambda}$. Then, $D(x, y) \leq \lambda$. Since $\lambda \leq \mu$, so $D(x, y) \leq \mu$. Hence $(x, y) \in \cong_{D,\mu}$.

(ii) Suppose that $(x, y) \in \cong_{D,\mu}$. Then $D(x, y) \leq \mu$. If $\lambda \geq \mu$, then it is obvious that$D(x, y) \leq \lambda$. Hence $(x, y) \in \cong_{D,\lambda}$ and we have $(x, y) \in \bigcap_{\lambda \geq \mu} \cong_{D,\lambda}$.

Conversely, if $(x, y) \in \bigcap_{\lambda \geq \mu} \cong_{D,\lambda}$ then $(x, y) \in \cong_{D,\mu}$ also holds. Therefore (a) is proved.

$b)$ To prove D is a dissimilarity, we should prove that it is reflexive, symmetric and transitive.

- Reflexive: Since $\cong_\lambda$ is an equivalence relation, for all $\lambda \in [0, 1]$, $x \cong_\lambda x$. Hence $D(x, x)$ = Inf $\{\lambda \in [0, 1] | x \cong_\lambda x\}$ =Inf [0,1] =0.
- Symmetric: Since Inf $\{\lambda \in [0, 1] | x \cong_\lambda y\}$ = Inf $\{\lambda \in [0, 1] | y \cong_\lambda x\}$, $D(x, y) = D(y, x)$.
- Transitivity: $D(x, z) \leq D(x, y) \vee D(y, z)$.

Set

$$A := \{\lambda \in [0, 1] | x \cong_\lambda y\},$$
$$B := \{\lambda \in [0, 1] | y \cong_\lambda z\},$$
$$C := \{\lambda \in [0, 1] | x \cong_\lambda z\}.$$

Let $\beta = D(x, z)$ and suppose that $\beta > D(x, y) \vee D(y, z)$, which implies that $\beta > D(x, y)$ and $\beta > D(y, z)$. By definition of infimum, $\beta > D(x, y)$ implies that there exists a $\lambda_1 \in [0, 1]$ such that $\beta > \lambda_1$ and $x \cong_{\lambda_1} y$. Also $\beta > D(y, z)$ implies that there exists a $\lambda_2 \in [0, 1]$ such that $\beta > \lambda_2$ and $y \cong_{\lambda_2} z$. Now, without loss of generality, let $\lambda_1 \leq \lambda_2$, then according to (i) of Proposition 3.4, $\cong_{\lambda_1} \subseteq \cong_{\lambda_2}$. That is, $x \cong_{\lambda_1} y$ implies that $(x, y) \in \cong_{\lambda_1}$ and because $\cong_{\lambda_1} \subseteq \cong_{\lambda_2}$, we have $(x, y) \in \cong_{\lambda_2}$ i.e. $x \cong_{\lambda_2} y$. Now since $x \cong_{\lambda_2} y$ and $y \cong_{\lambda_2} z$, according to the transitivity property of Definition 3.1, $x \cong_{\lambda_2} z$ and hence $\lambda_2 \in C$. Since β is InfC, we have $\beta \leq \lambda_2$ which contradicts $\beta > \lambda_2$. Therefore, $\beta \leq D(x, y) \vee D(y, z)$. Hence $D(x, z) \leq D(x, y) \vee D(y, z)$.

Now we prove that the family of $\lambda - cuts$ of dissimilarity is $\{\cong_\lambda\}_{\lambda \in [0,1]}$. For all $\lambda \in [0, 1]$, we know that $\cong_\lambda$ is an equivalence relation. Also $D(x, y)$ =Inf $\{\lambda \in [0, 1] | x \cong_\lambda y\}$. Therefore, D is a dissimilarity. Hence, $x \cong_{D,\lambda} y$ if and only if $D(x, y) \leq \lambda$. We prove that for all $\lambda \in [0, 1]$ the equality $\cong_\lambda = \cong_{D,\lambda}$ holds. Let $\mu \in [0, 1]$ be fixed and let $(x, y) \in \cong_\mu$. By definition of infimum, $D(x, y) \leq \mu$ and according to (1), we have $(x, y) \in \cong_{D,\mu}$. Therefore, $\cong_\mu \subseteq \cong_{D,\mu}$.
Conversly let$(x, y) \in \cong_{D,\mu}$ i.e. $x \cong_{D,\mu} y$. According to (1), $D(x, y) \leq \mu$.
We consider two cases:
Case 1: $D(x, y) < \mu$. That is, there exists $\lambda \in A$ that $\mu > \lambda \geq D(x, y)$. Then $x \cong_\lambda y$ and $\cong_\lambda \subseteq \cong_\mu$ which results $x \cong_\mu y$, hence $\cong_{D,\mu} \subseteq \cong_\mu$.
Case 2: $D(x, y) = \mu$. Let $\lambda > \mu$, then there exists $\lambda_1 \in A$ such that $\lambda > \lambda_1 \geq \mu$. Then

- $\lambda_1 \in A \Rightarrow x \cong_{\lambda 1} y$,
- $\lambda > \lambda_1 \Rightarrow \cong_{\lambda 1} \subseteq \cong_\lambda$.

It follows that $x \cong_\lambda y$.
We have proved that for all $\lambda > \mu$, we have $x \cong_\lambda y$. Hence, for all $n \in \mathbb{N}$, $\mu + \frac{1}{n} > \mu$ we have (x,y)$\in \cong_{\mu + \frac{1}{n}}$. On the other hand, by (ii), $\bigcap_{n=1}^{\infty} \cong_{\mu + \frac{1}{n}} = \cong_\mu$, therefore $(x, y) \in \cong_\mu$, and the proposition is proved.

Example 3.5: Let $U = \{M, B, G, E, P, H, F, T, C, W, O\}$, where:
$M = man, B = bear, G = gorilla, E = eagle, P = pigeon, H = hawk,$
$F = fish, T = tiger, C = cat, W = wolf, O = dog.$

We can define a dissimilarity D between elements in U by setting for any $x, y \in U$.
$D(x, y) = D(y, x)$
$D(x, y) = 0$ if $x = y$
$D(O, W) = D(E, H) = 0.36$

$D(M,G) = D(C,T) = 0.75$
$D(M,B) = D(B,G) = D(E,P) = D(H,P) = D(O,C) = D(O,T) = D(W,T) =$
$= D(W,C) = 0.96$
$D(x,y) = 1$ otherwise.

If we consider the relation $\cong_{D,0.4}$ $(\lambda = 0.4)$, the equivalence class of the element W is $\{W,O\}$ and the equivalence class of the element M is $\{M\}$. Again for the relation $\cong_{D,0.8}$, $(\lambda = 0.8)$, the equivalence class of the element W is $\{W,O\}$ and the equivalence class of the element M is $\{M,G\}$. An equivalent representation of D can be given by considering the quotient sets of the $\lambda_i - cuts$ in the family $\{\cong_{D,\lambda_i}\}_{0\le i\le 3}$, corresponding to the different similarity levels $\{\lambda_0,\lambda_1,\lambda_2,\lambda_3\} = \{0, 0.4, 0.8, 1\}$. We have

$U/\cong_{D,0} = \{\{M\},\{B\},\{G\},\{E\},\{P\},\{H\},\{F\},\{T\},\{C\},\{W\},\{O\}\}$,
$U/\cong_{D,0.4} = \{\{M\},\{B\},\{G\},\{E,H\},\{P\},\{F\},\{T\},\{C\},\{W,O\}\}$,
$U/\cong_{D,0.8} = \{\{M,G\},\{E,H\},\{F\},\{T,C\},\{W,O\},\{P\},\{B\}\}$,
$U/\cong_{D,1} = \{\{M,B,G,E,P,H,F,T,C,W,O\}\} = \{U\}$.

4. Relationship between similarity and dissimilarity relations

In many theoretical as well as practical applications, dissimilarity, rather than similarity, between two elements of a set needs to be considered. Especially when the elements are very similar to each other, studying the dissimilarities (differences) is easier. In this section we study functions that convert similarities to dissimilarities and vice versa. Rezaei and Mukaidono [21, 22] proved an equation that relates similarity and dissimilarity between two IFS's. Here we show that this equation can be used for elements of an ordinary set as well.

Proposition 4.1. *Let U be a domain and $D: U\times U \to [0,1]$ be a dissimilarity on U. Also let $f:[0,1]\to[0,1]$ be a decreasing function. Then $S: U\times U\to[0,1]$ defined by:*

$$S(x,y) = \frac{(f(D(x,y)) - f(1))}{f(0) - f(1)} \tag{2}$$

is a similarity on U.

Proof. It is enough to prove that S is a) Reflexive, b) Symmetric, and c) Transitive.
(a) Since $D(x,x) = 0$, for each $x\in U$ we have

$$S(x,x) = \frac{(f(D(x,x)) - f(1))}{f(0)-f(1)} = \frac{(f(0)) - f(1))}{f(0)-f(1)} = 1$$

(b) $S(x,y) = S(y,x)$, this condition follows since $D(x,y) = D(y,x)$ for each $(x,y)\in U$.
(c) We show that $S(x,z) \ge S(x,y)\wedge S(y,z)$.

$$S(x,z) = \frac{(f(D(x,z)) - f(1))}{f(0)-f(1)} \ge \frac{(f(D(x,y)\vee D(y,z)) - f(1))}{f(0)-f(1)}.$$

Since $D(x,z)\le D(x,y)\vee D(y,z)$, by Definition 3.1, we have $f(D(x,z)) \ge f(D(x,y))\vee f(D(y,z))$. Now, without loss of generality, suppose that $D(x,y)\le D(y,z)$. Then $f(D(x,y) \ge f(D(y,z))$. Consequently:

$$\frac{(f(D(x,z)) - f(1))}{f(0)-f(1)} \ge \frac{(f(D(x,y)) - f(1))}{f(0)-f(1)}.$$

Therefore, $S(x,z) \geq S(x,y) \wedge S(y,z)$.

Lemma 4.1. *Let $S : U \times U \to [0,1]$ be a similarity on U. Also let $f : [0,1] \to [0,1]$ be a decreasing function. Then $D : U \times U \to [0,1]$ defined by:*

$$D(x,y) = \frac{(f(S(x,y)) - f(1))}{f(0) - f(1)} \tag{3}$$

is a dissimilarity on U.

Proof. The proof is similar to that of Proposition 4.1

4.1. **Convex and Concave functions.** We recall some definitions pertaining to concave and convex functions.

Definition 4.1. *A real-valued function $f : C \longrightarrow R$ defined on an interval or on any convex set C of a vector space is called concave, if for any two points x and y in its domain C and any t in [0,1], we have*

$$f(tx + (1-t)y) \geq tf(x) + (1-t)f(y). \tag{4}$$

Similarly f is called convex if for any two points x and y in its domain C and any t in [0,1], we have

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y). \tag{5}$$

f is called strictly convex if

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$

for any t in [0,1] and $x \neq y$.

In particular, f is concave on $[a,b]$ if and only if the function $-f$ is convex on $[a,b]$.
We know that information about two different elements in any universe of discourse is not more than the overall information between them and we expect the same is true for their similarities and dissimilarities. That is, the information regarding the similarities and dissimilarities of two members of any set should be enough to draw any conclusion about them. So, when we study different sorts of sets, we expect for any members like x and y of any set, the following relation holds:

$$D(x,y) + S(x,y) \leq 1. \tag{6}$$

The decreasing functions that relate similarities and dissimilarities, like function f in (2) and (3), can be convex or concave. In the following proposition we prove that if the function is convex, then (6) holds.

Proposition 4.2. *Let $f : [0,1] \to [0,1]$ be a decreasing function and D be a dissimilarity on U and the relation*

$$S(x,y) = \frac{(f(D(x,y)) - f(1))}{f(0) - f(1)}$$

holds. Then we have:
(a) $D(x,y) + S(x,y) \geq 1$ if f is concave,
(b) $D(x,y) + S(x,y) = 1$ if f is the function $f(t) = 1 - t$ for $t \in [0,1]$,
(c) $D(x,y) + S(x,y) \leq 1$ if f is convex.
Proof. *We prove (a). The proof of (b) and (c) is similar. We know that $0 \leq D(x,y) \leq 1$. Now if $0 < D(x,y) < 1$ and f is concave, we have:*

$$\frac{f(D(x,y)) - f(1)}{D(x,y) - 1} \leq \frac{f(0) - f(1)}{0 - 1}.$$

That is,

$$S(x,y) = \frac{f(D(x,y)) - f(1)}{f(0) - f(1)} \geq \frac{D(x,y) - 1}{0 - 1} = 1 - D(x,y)$$

Hence

$$D(x,y) + S(x,y) \geq 1.$$

Remark 1: Let $f : [0,1] \to [0,1]$ be a decreasing function, S be a similarity on a fuzzy set U and D be the dissimilarity defined in (3), then the conclusion of Proposition 4.4 is true.

We close this section by an example which reveals the ideas proposed in this paper.

Example 4.5 : Using the of Example 3.5, we can define a similarity S between elements in U by setting for any $x, y \in U$

$S(x,y) = S(y,x)$
$S(x,y) = 1$ if $x = y$
$S(O,W) = S(E,H) = 0.8$
$S(M,G) = S(C,T) = 0.5$
$S(M,B) = S(B,G) = S(E,P) = S(H,P) = S(O,C) = S(O,T) = S(W,T) = S(W,C) = 0.2$
$S(x,y) = 0$ otherwise.
Also consider the functions:

$$f_1(x) = 1 - x$$
$$f_2(x) = 1 - \sqrt{x}$$
$$f_3(x) = (1 - \tanh(x)) - \alpha x \quad \mathit{for\ some}\ \alpha > 0.$$

We select $\alpha = 0.24$ to keep the values in the range [0,1]. f_2 and f_3 are convex.
Now using the Remark 1 and the relation (3), from f_1, f_2, and f_3 we obtain dissimilarities D_1, D_2, and D_3 respectively (Table 1).

TABLE 1. The level of similarity of two elements and computed dissimilarities according to the different functions.

The elements	$\mathcal{S}$	$\mathcal{D}1$	$\mathcal{D}2$	$\mathcal{D}3$
(M,F)	0	1	1	1
(M,B)	0.2	0.8	0.55	0.76
(C,T)	0.5	0.5	0.29	0.42
(O,W)	0.8	0.2	0.1	0.15
(F,F)	1	0	0	0

We note that the relation

$$D_1(x,y) + S(x,y) = 1$$

holds for f_1, and the inequality

$$D_i(x,y) + S(x,y) \leq 1 \quad \mathit{for}\ i = 2, 3$$

holds for convex functions f_2 and f_3.
Now using the information in Table 1, we compute the sum of similarities and dissimilarities computed using the functions f_1, f_2, and f_3. The results are in Table 2.
The dissimilarity for the concave function $f_4(x) = 1 - x^2$ is shown in Table 3.

TABLE 2. The sum of similarity and dissimilarity

$\mathcal{S}$	$\mathcal{S}+\mathcal{D}1$	$\mathcal{S}+\mathcal{D}2$	$\mathcal{S}+\mathcal{D}3$
0	1	1	1
0.2	1	0.75	0.96
0.5	1	0.79	0.92
0.8	1	0.9	0.95
1	1	1	1

TABLE 3. The level of similarity and computed dissimilarity and their sum for f_4.

The elements	$\mathcal{S}$	$\mathcal{D}4$	$\mathcal{S}+\mathcal{D}4$
(M,F)	0	1	1
(M,B)	0.2	0.96	1.16
(C,T)	0.5	0.75	1.25
(O,W)	0.8	0.36	1.16
(F,F)	1	0	1

As it can be seen $S(x,y)+D_4(x,y)\geq 1$. This example shows that we can convert a similarity to dissimilarity if we use a decreasing convex function. Note that D_4 is the dissimilarity which was used in Example 3.5.

5. Similarity and dissimilarity relations in a finite universe

In Propositions 4.1 and 4.2, we tried to obtain similarity (dissimilarity) from dissimilarity (similarity) utilizing a decreasing function. Convexity (or concavity) of this function helped us to find the relationship between similarity and dissimilarity measures. Now the question arises: If there is a similarity and a dissimilarity, can we always find a decreasing function that can be used to relate similarity and dissimilarity as in (2)?

If the universe U is a finite set, then a similarity relation S and a dissimilarity relation D have been defined on $U\times U$. We introduce a decreasing function $f:[0,1]\to[0,1]$ that can set up a linear (or partially linear) relation between the similarity and dissimilarity relations. The following proposition gives the answer for a finite universe.

Proposition 5.1. *Let $U\times U=\{(x_1,y_1),....,(x_n,y_n)\}$ for some positive integer n. Suppose that S is a similarity and D is a dissimilarity relation on U such that $s(x_i,y_i)=0$ only if $D(x_i,y_i)=1$. Then there is a continuous decreasing function $f:[0,1]\to[0,1]$ and positive integer numbers k_i such that*

$$f(D(x_i,y_i))=\frac{1}{k_i}S(x_i,y_i) \qquad for\ \ i=1,2,3,...,n. \quad (7)$$

Proof. Without loss of generality let $D(x_1,y_1)<D(x_2,y_2)<\ldots<D(x_n,y_n)$. Set $k_1=1$ and compute $k_2,k_3,\ldots,k_n$ as follows:

$$k_{i+1}=\begin{cases}1 & \text{if } S(x_i,y_i)\text{=0;}\\ [\frac{k_iS(x_{i+1},y_{i+1})}{S(x_i,y_i)}]+1 & \text{otherwise,}\end{cases}$$

where [x] is the greatest integer number less than or equal to x. Now define $f(D(x_i,y_i))=\frac{1}{k_i}S(x_i,y_i)$ for i=1,2,3,....n. Then, we have to show

$$f(D(x_{i+1},y_{i+1}))\leq f(D(x_i,y_i)) \qquad for\ \ i=1,2,3,...,n-1.$$

If $S(x_i, y_i) = 0$, then $D(x_i, y_i) = 1$ and $i = n$, because $i < n$ implies that

$$D(x_{i+1}, y_{i+1}) > D(x_i, y_i) = 1,$$

that is

$$D(x_{i+1}, y_{i+1}) > 1$$

which is a contradiction.
Now let $1 \leq i \leq n$ and $S(x_i, y_i) \neq 0$. We have

$$[\frac{k_i S(x_{i+1}, y_{i+1})}{S(x_i, y_i)}] \leq \frac{k_i S(x_{i+1}, y_{i+1})}{S(x_i, y_i)} < [\frac{k_i S(x_{i+1}, y_{i+1})}{S(x_i, y_i)}] + 1 = k_{i+1}.$$

That is

$$\frac{k_i S(x_{i+1}, y_{i+1})}{S(x_i, y_i)} < k_{i+1}.$$

Then

$$\frac{1}{k_{i+1}} S(x_{i+1}, y_{i+1}) < \frac{1}{k_i} S(x_i, y_i).$$

Therefore

$$f(D(x_{i+1}, y_{i+1})) \leq f(D(x_i, y_i)).$$

So far we have defined f on the set $\{D(x_1, y_1), D(x_2, y_2)......., D(x_n, y_n)\}$ and we can extend f on intervals $(D(x_i, y_i), D(x_{i+1}, y_{i+1}))$ as a line that connects $(D(x_i, y_i), f(D(x_i, y_i)))$ and $(D(x_{i+1}, y_{i+1}), f(D(x_{i+1}, y_{i+1})))$. Hence f is a continuous decreasing function on [0,1].

As it can be seen $k_i' s$ are positive integers generated using S. The role of k_i's in (7) resembles the role of $f(0) - f(1)$ in (2). It should be mentioned that there are other ways to obtain k_i's as well. This fact stresses that the obtained function in Proposition 5.1 is not unique.

Example 5.2: We have developed an algorithm (Appendix 1) to implement Proposition 5.1. In this algorithm we first generate two sets of 30 random numbers in the range [0,1] and place them into two arrays representing corresponding dissimilarity D and similarity S values. After sorting the dissimilarity values in ascending order, we find k_i's, the coefficients explained in Proposition 5.1, and the values of $f(D(x_i, y_i))$. Table 4 shows the results of our computations. According to Proposition 5.1 and using the algorithm, we have obtained the decreasing function that relates *similarity* and *dissimilarity*. As it can be seen in Figure 1, the function initially decreases steeply, then it approaches a constant function.

6. Conclusion

A set U can be divided into different equivalence classes of dissimilarities at different levels. In some cases we might obtain dissimilarities from similarities using the relation $D(x, y) = 1 - S(x, y)$. However, there are other decreasing functions that may model the relation between similarity and dissimilarity. In the real world we expect that the sum of a similarity and a dissimilarity does not exceed 1 and this is possible when the function is convex. It is obvious that choosing a proper function is very important. When similarity and dissimilarity are known, the existence of a continuous decreasing function to relate them is essential. Now the question is: What is the simplest function for this purpose? Can it be a smooth or even a linear function? Future works could consist of generalizing dissimilarities using T-norms or T-conorms. For example we may use a T-norm instead of minimum in the third condition of the similarity definition (Definition 2.1) to get T-similarity. In the same way we can get T-dissimilarity. Now the question is that whether there are continuous decreasing functions that relate T-similarity and T-dissimilarity.

Table 4. Dissimilarities, Similarities, Ki's, and FD's of Example 5.2

$\mathcal{D}$	$\mathcal{S}$	$\mathcal{K}i$	$\mathcal{FD}$
0	1	1	1
0.025736	0.730164	1	0.730164
0.160278	0.110279	1	0.110279
0.314896	0.838092	8	0.104761
0.315109	0.603765	6	0.100627
0.339649	0.637622	7	0.091089
0.392882	0.089939	1	0.089939
0.470649	0.46114	6	0.076857
0.512402	0.803775	11	0.07307
0.566672	0.272491	4	0.068123
0.583303	0.321667	5	0.064333
0.602995	0.550327	9	0.061147
0.6123	0.983063	17	0.057827
0.652122	0.602842	11	0.054804
0.666726	0.958378	18	0.053243
0.669273	0.243246	5	0.048649
0.692717	0.199505	5	0.039901
0.696766	0.699108	18	0.038839
0.717231	0.49293	13	0.037918
0.751351	0.565893	15	0.037726
0.787572	0.216347	6	0.036058
0.848359	0.29909	9	0.033232
0.863388	0.763774	23	0.033208
0.88624	0.419153	13	0.032243
0.898677	0.91379	29	0.03151
0.912537	0.719493	23	0.031282
0.912896	0.542239	18	0.030124
0.959565	0.929327	31	0.029978
0.986567	0.226258	8	0.028282
1	0	1	0

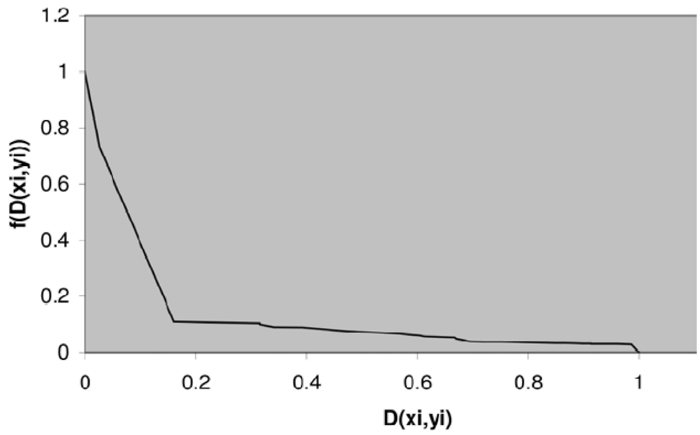

Figure 1. The relationship between similarity and dissimilarity of Example 5.2.

7. Appendix 1

Algorithm 1

(1) Set SIZE=30
 /* SIZE is the number of dissimilarity and similarity values,
 30 in this example. */

(2) Set dis[SIZE]=0, sim[SIZE]=0
 /* Set arrays *dis* and *sim* of SIZE elements to hold dissimilarity and
 and similarity values initially to zero.*/

(3) FOR i=1 to SIZE /*Fill arrays dis and sim with
 generated random numbers between 0 and 1.*/
 Set dis[i]=rand() /* rand() is random nuber generator function*/
 Set sim[i]=rand()
 ENDFOR

(4) Sort(dis) /* Sort array *dis* in ascending order to meet the condition:
 $D(x_1, y_1) < D(x_2, y_2) < \ldots < D(x_n, y_n)$.
 omit the identical values,
 rearrange corresponding *sim* array accordingly.*/

(5) Set k[1]=1 /* set $k_1 = 1$ */

(6) FOR i=1 to SIZE /* compute $k_2, k_3, \ldots k_i$*/
 IF sim[i]=0 THEN /* if $S(x_i, y_i)$=0 */
 Set k[i+1]=1 /* $k_{i+1} = 1$ */
 ELSE
 Compute k[i+1] $= [\frac{k_i * sim[i+1]}{sim[i]}] + 1$
 ENDIF
 ENDFOR

(7) /* Compute $f(D(x_i, y_i))$'s. */
FOR i=1 to SIZE /* compute $f(D(x_i, y_i))$,i=1,2,3,....,SIZE */
Compute f[i]= $\frac{1}{k_i} * sim[i]$ /* $f(D(x_i, y_i))=\frac{1}{k_i} * S(x_i, y_i)$ */
ENDFOR
(8) END /* of algorithm */

Acknowledgements

The authors would like to thank professor M. Tata for her careful editing of the paper. The third author is partially supported by Fuzzy Systems and Applications Center of Excellence, Shahid Bahonar University of Kerman.

References

[1] Alguliev, R., Aliguliyev, R. Eperimental investigating the F-Meaure as similarity measure for automatic text summarization, Appl. and Comput. Math., V.6, N.2, 2007, pp.278-287.
[2] Aliev, R. A. Modelling and stability analysis in fuzzy economics, Appl. and Comput. Math., V.7, N.1, 2008, pp.31-53.
[3] Atanassov, K. T. Intuitionistic fuzzy set, Fuzzy Sets and Systems, V.20, 1986, pp.87-96.
[4] Atanassov, K. T. Intuitionistic fuzzy sets, theory and applications, Studies in Fuzziness and Soft Computing, V.35, 1999, Physica-Verlag, Heidelberg.
[5] Bisson, G. Why and how to define a similarity measure for object-based representation systems, In N.J.I. Mars, editor, Towards Very Large Knowledege Bases, 1992, pp.236-246.
[6] Chakrabotry, M., Das, S. On fuzzy equivalence, Fuzzy Sets and Systems, V.11, 1983, pp.185-193.
[7] Chen, S. M. A new approach to handling fuzzy decision making problems, IEEE Trans. Systems, Man. Cybernet, V.18, 1988, pp.1012-1016.
[8] Deschrijver, G., Kerr, E.E. On the position of intuitionistic fuzzy set theory in the framework of theories modelling impercision, Information Sciences, V.177, 2007, pp.1860-1866.
[9] Goldstone, R., Son, J. Similarity, in cambridge handbook of thinking and reasoning Cambridge University Press: Cambridge, 2005.
[10] Goldstone, R., Medin, D., and Gentner, D. Respects for similarity, Psychological Review, V.100, N.2, 1993, pp.254-278.
[11] Jacas, J., Valverde, L. On fuzzy relations, metric and cluster analysis, In: Delgado M, Verdegay, JL, (Eds) Approximate reasoning Tools for Artificial Intelligencem Rheinland: Verlag TV, 1990, pp. 204-220.
[12] Jain, A., Dubes, R. Algorithm for clustring data, Prentice Hall, 1988.
[13] Klawonn, F., Castro, J. L. Similarity in Fuzzy reasoning, Mathware & Soft computing, V.3, 1995, pp.197-228.
[14] Loia, V., Senatore, S. Sessa, M., Combining agent technology and similarity-based reasoning for targeted E-mail services, Fuzzy Sets and Systems, V.145, 2004, pp. 29-56.
[15] Nikolova, M., Nikolov, N., Cornelis, C., Deschrijver, G. Survey of the research on intuitionistic fuzzy sets, Advanced Studies in Contemporary Mathematics, V.4, N.2, 2002, pp.127-157.
[16] Ovchinnikov, S. V. Structure of fuzzy binary relations, Fuzzy Sets and Systems, V.6, 1981, pp.169-195.
[17] Ovchinnikov, S. V. Representations of transitive fuzzy relations, In: Skala HJ et al. (Eds) Aspects of Vagueness, Dordrecht: Reidel, 1984, pp.105-118.
[18] Pappis, C. P., Karacapilidis, N. I. A comarative assessment of measures os similarity of fuzzy values, Fuzzy Sets and Systems, V.56, 1993, pp.171-174.
[19] Rezaei, H., Mukaidono, M. Some results on similarity measures of intuitionistic fuzzy sets, Proceedings of AFSS, V.2, 2006, pp.112-120.
[20] Rezaei, H., Mukaidono, M. New similarity measures of intuitionistic fuzzy sets, Journal of Advanced Computational Intelligence and Intelligent Informatics, V.11, N.2, 2007, pp. 202-209.

[21] Rezaei, H., Emoto, M., Mukaidono, M. New similarity measure between two fuzzy sets, Journal of Advanced Computational Intelligence and Intelligent Informatics, V.10, N.6, 2006, pp.946-953 .

[22] Rezaei, H. Fuzzy Logic and its applications in artificial intelligence: Intiuition-based similarity measures, Ph.D. Thesis, Department of Computer Science, Meiji University, 2006.

[23] Rissland, E. L. AI and similarity, IEEE Intelligent Systems, V.21, N.3, 2006, pp.39-49.

[24] Rodriguez, A., Egenhofer, M. Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure, International Journal of Geographical Information Science, V.18, N.3, 2004, pp.229-256.

[25] Trillas, E., Valverde, L. An inquiry on indistinguishability operators, In: Skala HJ et al. (Eds) Aspects of Vagueness, Dordrecht: Reidel, 1984, pp.512-581.

[26] Valtchev, P. Construction automatique de taxanomies pour l'aide a' la representaion de connaissance par objects, These de doctorat, Universite Joseph Fourier, Grenoble I, 1999.

[27] Valverde, L. On the structure of F-indistinguishability operators, Fuzzy Sets and Systems, V. 17, 1995 pp.313-328.

[28] Zadeh, L. A. Fuzzy sets. Information Control, 1965, pp.338-353.

[29] Zwick, R., Carlstein, E., and Budesco, D. Measures on fuzzy sets: A comparative analysis, Int. J. Approximate reasoning, V.1, 1987, pp.221-242.

Mahmood Keshavarzi - He was born in 1954 in Rafsanjan, Iran. He received his B.Sc. (1971) in Statistics from The Institute of Statistics and Informatics in Tehran, Iran. His M.Sc. (1980) Computer Science from The University of Southern Mississippi, Hattiesburg, Mississippi, U.S.A. Currently he is a lecturer in the department of computer science in Vali Asr University of Rafsanjan, Iran. At the same time he is working on his Ph.D. at the Shahid Bahonar University of Kerman (SBUK) in Iran. His current interest is in fuzzy mathematics, especially its applications on soft computing.

Mohammad Ali Dehghan - He was born in 1962 in Rafsanja,Iran. He received his B.Sc. (1985), his M.Sc. (1989), and his Ph. D. (1994) all in Mathematics from Shahid Bahonar University of Kerman, Iran. He is an associate professor in the department of Mathematics at the Vali Asr University of Rafsanjan in Iran. His current interest is in wavelets and frames and functional analysis.

Mashaallah Mashinchi, for a photograph and biography, see Appl. and Comput. Math., V.7, N.1, 2008, p.125.