Fifteenth Int. Conf. on IFSs, Burgas, 11-12 May 2011
NIFS 17 (2011), 2, 68-74

A note on distance and similarity measures between intuitionistic fuzzy sets

Peter Vassilev

IBPhBME – Bulgarian Academy of Science
105 Acad. G. Bonchev Str, 1113 Sofia, Bulgaria
Bulgarian Academy of Sciences
E-mail: peter.vassilev@gmail.com

Abstract: In the present paper some of the proposed in the literature distances and pseudometrics between intuitionistic fuzzy sets are considered and a more general approach for constructing new distances, with their help, is proposed. As a result, a way of generating new similarity measures is given.
Keywords: Intuitionistic fuzzy sets, Distance, Metric, Pseudometric, Similarity measure.
AMS Classification: 03E72

1 Introduction

Intuitionistic fuzzy sets (IFS) were introduced as a generalization and extension of the concept of Fuzzy sets, by K. Atanassov (see [1]). We will briefly remind some of the basic definitions and notions.

Let X be a universe set (i.e. the set of all the (relevant) elements that will be considered) and let $A \subset X$. An intuitionistic fuzzy set is a set

$$A^* \stackrel{\text{def}}{=} \{\langle x, \mu_A(x), \nu_A(x)\rangle | x \in X\}$$

where the functions

$$\mu_A : X \to [0,1]$$

and

$$\nu_A : X \to [0,1]$$

reflect the degree of membership and non-membership of the element x from X to the set A, respectively, and for every $x \in X$ it is fulfilled that:

$$0 \le \mu_A(x) + \nu_A(x) \le 1 \tag{1}$$

The function $\pi_A : X \to [0,1]$, which is given by

$$\pi_A(x) \stackrel{\text{def}}{=} 1 - \mu_A(x) - \nu_A(x)\ \forall x \in X \tag{2}$$

defines the degree of uncertainty of the membership of the element x to the set A.

Many distances and pseudometrics have been defined over discrete IFS (the most common on practice). They are usually introduced as (weighted) sum of the distances between their respective points.

First, we will start with the formal definitions (see e.g. [2]).

Definition 1. *Let A, B and C be discrete IFS (defined over the same universe X). The class of all such sets is further denoted by IFS(X). Then, the distance is a function* $d : IFS(X) \times IFS(X) \to [0, +\infty)$ *such that:*

1. $d(A, B) = 0$ *if and only if* $A = B$
2. $d(A, B) = d(B, A)$ *(symmetry)*
3. $d(A, B) + d(B, C) \geq d(A, C)$ *(triangle inequality)*

Remark 1. *When condition 1. from Definition 1 is relaxed, i.e.* $d(A, B) = 0$ *not only when* $A = B$, d *is called a pseudometric. In practice, it is usually ensured that* $d : IFS(X) \times IFS(X) \to [0, 1]$, *since it can provide a nice relation between similarity measures defined over IFS and appropriate distances.*

Let us briefly remind two of the most commonly used distances in unweighted form for IFSs defined over a discrete universe $X = \{x_1, \ldots, x_n\}$:

$$d(A, B) = \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| \tag{3}$$

(see [1])

$$d(A, B) = \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)| \tag{4}$$

(see [3]).

Here we will introduce infinitely many pseudometrics, namely:

$$\varrho_k(A, B) = \sum_{i=1}^{n} |(\mu_A^k(x_i) - \nu_A^k(x_i)) - (\mu_B^k(x_i) - \nu_B^k(x_i))|,$$

where $k = 1, 2, \ldots$

Another pseudometric is given in [4]. Namely,

$$\varrho^*(A, B) = \arccos\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\mu_A(x_i)\mu_B(x_i) + \nu_A(x_i)\nu_B(x_i)}{\sqrt{\mu_A^2(x_i) + \nu_A^2(x_i)}\sqrt{\mu_B^2(x_i) + \nu_B^2(x_i)}}\right) \tag{5}$$

For the case $n = 1$ it is seen that if $\varrho^*(A, B) = 0$ then we have

$$\mu_A(x_1)\mu_B(x_1) + \nu_A(x_1)\nu_B(x_1) = \sqrt{\mu_A^2(x_1) + \nu_A^2(x_1)}\sqrt{\mu_B^2(x_1) + \nu_B^2(x_1)}$$

But, because of the fact, that the well known Cauchy's inequality

$$\mu_A(x_1)\mu_B(x_1) + \nu_A(x_1)\nu_B(x_1) \leq \sqrt{\mu_A^2(x_1) + \nu_A^2(x_1)}\sqrt{\mu_B^2(x_1) + \nu_B^2(x_1)}$$

becomes equality if and only if (for our case):

$$\frac{\mu_A(x_1)}{\mu_B(x_1)} = \frac{\nu_A(x_1)}{\nu_B(x_1)} \tag{6}$$

it is clear that the fulfillment of:

$$\mu_A = \mu_B,\ \nu_A = \nu_B$$

is not a necessary condition for the validity of (6). Thus $\varrho^*(A, B)$ is a pseudometric and not a distance.

It can be directly checked that for any of the above distances d (like (3) and (4)) we have:

$$d^*(A, B) \stackrel{\text{def}}{=} d(A, B) + \varrho(A, B)$$

(where ϱ is an arbitrary pseudometric) is a distance. Moreover, this is true for an arbitrary distance d and pseudometric ϱ.

Definition 2. *Let A, B and C be discrete IFS (defined over the same universe X). Then, the similarity measure is a function* $s : IFS(X) \times IFS(X) \to [0, 1]$ *such that*

1. $s(A, B) = 1$ *if and only if* $A = B$
2. $s(A, B) = s(B, A)$
3. *If* $A \subseteq B \subseteq C$*, then:*

$$s(A, C) \leq s(A, B);\ s(A, C) \leq s(B, C)$$

Remark 2. *If*

$$s(A, B) = 1$$

but $A \neq B$ *we will call* s *pseudo-similarity measure.*

Remark 3. *As it can easily seen, there is a nice relation between appropriate (weighted) distances and similarity measures, namely:*

$$s(A, B) = 1 - d(A, B) \tag{7}$$

Different types of distances and similarity measures reflect well different types of useful information. Therefore, ways to produce distance (similarity measure) which is more informative are desirable.

2 Generating new distances

Some of the distances and pseudometrics in the previous section bear much information in some cases and less in others. Here, we will propose a more general approach at producing more suitable distances and similarity measures.

First we need the following

Lemma 1. *Let* $f : IFS(X) \times IFS(X) \to [0, +\infty)$ *be a pseudometric (metric). Then for arbitrary real number* $\alpha \in (0, 1]$ *the mapping which is given by*

$$f_\alpha(A, B) \stackrel{\text{def}}{=} (f(A, B))^\alpha \ \ \forall A, B \in IFS(X) \tag{8}$$

is a pseudometric (metric) too.

Proof. To prove that f_α is a pseudometric (metric), we need only to establish that for f_α the triangle inequality holds. For that aim it is enough to verify that for every three IFSs A, B and C the inequality

$$f_\alpha(A,B) + f_\alpha(B,C) \geq (f(A,B) + f(B,C))^\alpha$$

holds.

But the above inequality is equivalent to the inequality

$$\left(\frac{f(A,B)}{f(A,B)+f(B,C)}\right)^\alpha + \left(\frac{f(B,C)}{f(A,B)+f(B,C)}\right)^\alpha \geq 1$$

The last inequality is certainly true, since we may put:

$$u = \frac{f(A,B)}{f(A,B)+f(B,C)}; v = \frac{f(B,C)}{f(A,B)+f(B,C)}$$

and observe that $u + v = 1$ and then $u^\alpha + v^\alpha \geq 1$ because of the fact that $\alpha \in (0,1]$.

The Lemma is proved. □

The following proposition is obvious:

Proposition 1. *Let $A, B, C \in IFS(X)$ are such that $A \subseteq B \subseteq C$ and let $\alpha \in (0,1]$ be an arbitrary real number. If $f : IFS(X) \times IFS(X) \to [0, +\infty)$ is a pseudometric (metric) and the inequalities:*

$$f(A,B) \leq f(A,C) \; f(B,C) \leq f(A,C)$$

hold, then f_α (given by (8)) is also a pseudometric (metric) and the inequalities

$$f_\alpha(A,B) \leq f_\alpha(A,C) \, ; f_\alpha(B,C) \leq f_\alpha(A,C)$$

hold.

As a corollary of Lemma 1 and Proposition 1 we obtain the following general result.

Theorem 1. *Let n and k be arbitrary positive integers. Let $f^{(i)} : IFS(X) \times IFS(X) \to [0, +\infty), i = \overline{1,n}$ be metrics and $g^{(j)} : IFS(X) \times IFS(X) \to [0, +\infty), j = \overline{1,k}$ be pseudometrics. Then, for every pair of fixed real numbers $\alpha_i \in (0,1], i = \overline{1,n}$ and $\beta_j \in (0,1], j = \overline{1,k}$, the mapping $t_{\alpha,\beta} : IFS(X) \times IFS(X) \to [0, +\infty)$ that is given by:*

$$t_{\alpha,\beta}(A,B) = \sum_{i=1}^{n} f_{\alpha_i}^{(i)}(A,B) + \sum_{j=1}^{k} g_{\beta_j}^{(j)}(A,B), \tag{9}$$

where $\alpha \stackrel{\text{def}}{=} (\alpha_1, \ldots, \alpha_n)$, $\beta \stackrel{\text{def}}{=} (\beta_1, \ldots, \beta_k)$ is a metric.

If $A, B, C \in IFS(X)$ are such that $A \subseteq B \subseteq C$ and we have:

$$f^{(i)}(A,B) \leq f_i(A,C); \; f^{(i)}(B,C) \leq f^{(i)}(A,C), \; \overline{i = 1,n};$$

$$g^{(j)}(A,B) \leq g^{(j)}(A,C); \; g^{(j)}(B,C) \leq g^{(j)}(A,C), \; \overline{j = 1,k},$$

then it is true:

$$t_{\alpha,\beta}(A,B) \leq t_{\alpha,\beta}(A,C); \; t_{\alpha,\beta}(B,C) \leq t_{\alpha,\beta}(A,C)$$

Further, we will solve the problem how to transform a given metric $d : IFS(X)\times IFS(X) \to [0,+\infty)$ into a normalized metric d^* i.e. $d^* : IFS(X)\times IFS(X) \to [0,1]$. This problem is solved by the following

Theorem 2. *Let $\xi > 0$ and $\eta \geq 1$ be fixed real numbers and $d : IFS(X) \times IFS(X) \to [0,+\infty)$ be an arbitrary metric (pseudometric). Then, the mapping $d^* : IFS(X) \times IFS(X) \to [0,1]$ that is given by*

$$d^*(A,B) \overset{\text{def}}{=} \frac{d(A,B)}{\xi + \eta d(A,B)} \tag{10}$$

is a normalized metric (pseudometric). Moreover, if d is a non-Archimedean metric, then d^* is a non-Archimedean metric, too.

Proof. To prove the fact that d^* is a metric (pseudometric), we need only to prove that the triangle inequality for d^* holds. This means that for every three IFSs A, B and C, the inequality

$$\frac{d(A,B)}{\xi + \eta d(A,B)} + \frac{d(B,C)}{\xi + \eta d(B,C)} \geq \frac{d(A,C)}{\xi + \eta d(A,C)}$$

holds.

First let $d(A,C) \geq \max(d(A,B), d(B,C))$. Then, we have:

$$\frac{d(A,B)}{\xi + \eta d(A,B)} + \frac{d(B,C)}{\xi + \eta d(B,C)} \geq \frac{d(A,B)}{\xi + \eta d(A,C)} + \frac{d(B,C)}{\xi + \eta d(A,C)} \geq \frac{d(A,B) + d(B,C)}{\xi + \eta d(A,C)}$$
$$\geq \frac{d(A,C)}{\xi + \eta d(A,C)},$$

since the triangle inequality is valid for d.

Second, let $d(A,C) \leq \max(d(A,B), d(B,C))$. Then, supposing without loss of generality, that $\max(d(A,B), d(B,C)) = d(B,C)$, we consider the only possible cases:

1) $d(A,C) = 0$;

2) $d(A,C) > 0$

In case 1) the triangle inequality for d^* is obviously fulfilled. Let case 2) hold. Then, using that $0 < d(A,C) \leq d(B,C)$, we have

$$\frac{d(A,C)}{\xi + \eta d(A,C)} = \frac{1}{\eta + \frac{\xi}{d(A,C)}} \leq \frac{1}{\eta + \frac{\xi}{d(B,C)}} = \frac{d(B,C)}{\xi + \eta d(B,C)} \leq \frac{d(A,B)}{\xi + \eta d(A,B)} + \frac{d(B,C)}{\xi + \eta d(B,C)}$$

Thus, the triangle inequality for d^* is proved and therefore d^* is a metric (pseudometric). From (10) it is seen that $d^*(A,B) < 1$, i.e. d^* is a normalized metric (pseudometric).

It remains only that if d is a non-Archimedean metric, d^* is a non-Archimedean metric, too. Let us suppose that d is a non-Archimedean metric. This means that for every three IFSs A, B and C (defined over the universe X), it is fulfilled:

$$d(A,B) \leq \max(d(B,C), d(A,C)). \tag{11}$$

Then, to prove that d^* is also a non-Archimedean metric, we must verify the validity of the inequality

$$\frac{d(A,B)}{\xi + \eta d(A,B)} \leq \max\left(\frac{d(B,C)}{\xi + \eta d(B,C)}, \frac{d(A,C)}{\xi + \eta d(A,C)}\right)$$

If $d(A,B)=0$, the above inequality is obvious. Let $d(A,B)>0$. Without loss of generality, let $\max(d(A,C),d(B,C))=d(B,C)$. Then, since (11) is valid, we have

$$d(A,B)\le d(B,C).$$

Hence:

$$d^*(A,B)=\frac{d(A,B)}{\xi+\eta d(A,B)}=\frac{1}{\eta+\frac{\xi}{d(A,B)}}\le\frac{1}{\eta+\frac{\xi}{d(B,C)}}=\frac{d(B,C)}{\xi+\eta d(B,C)}$$

$$\le\max\left(\frac{d(B,C)}{\xi+\eta d(B,C)},\frac{d(A,C)}{\xi+\eta d(A,C)}\right)=\max\left(d^*(B,C),d^*(A,C)\right)$$

The Theorem is proved. □

Corollary 1. *Let $A,B,C\in IFS(X)$ are such that $A\subseteq B\subseteq C$. If $d:IFS(X)\times IFS(X)\to[0,+\infty)$ is an arbitrary metric for which it is fulfilled*

$$d(A,B)\le d(A,C);\ d(B,C)\le d(A,C),$$

and d^* is the metric that is given by (10)*, then the above inequalities are valid with d^* instead of d.*

Proof. We will only prove that $d^*(A,B)\le d^*(A,C)$, since the second inequality may be proved in the same manner. Assuming $d(A,B)>0$ we have (because of (10) and $d(A,B)\le d(A,C)$) :

$$d^*(A,B)=\frac{d(A,B)}{\xi+\eta d(A,B)}=\frac{1}{\eta+\frac{\xi}{d(A,B)}}\le\frac{1}{\eta+\frac{\xi}{d(A,C)}}=\frac{d(A,C)}{\xi+\eta d(A,C)}=d^*(A,C)$$

and the corollary is proved. □

Below we need the following proposition

Proposition 2. *Let $n>1$ be a positive integer and $d^{(i)},i=\overline{1,n}$ are different mappings $d^{(i)}:IFS(X)\times IFS(X)\to[0,+\infty)$ such that each one of them is a metric or pseudometric and at least one of them is a metric. Then, an arbitrary linear combination of these mappings with positive coefficients represents a metric d over $IFS(X)\times IFS(X)$.*

Remark 4. *The above mentioned metric d may be transformed by infinitely many ways (changing ξ and η) into normalized metric d^* that is given by* (10)*.*

Using the mentioned in Proposition 2 and Remark 4, from Theorem 1, we obtain a way of introducing **infinitely many new similarity measures**. For this aim we use the infinitely many metrics $t_{\alpha,\beta}$ (see (9)). For each of them, we introduce the normalized metric $t^*_{\alpha,\beta}$ that is given by

$$t^*_{\alpha,\beta}(A,B)=\frac{t_{\alpha,\beta}(A,B)}{\xi+\eta t_{\alpha,\beta}(A,B)}\ \ \forall A,B\in IFS(X)$$

with arbitrary real numbers $\xi>0,\ \eta\ge 1$ and having the property:

$$t^*_{\alpha,\beta}(A,B)\le t^*_{\alpha,\beta}(A,C);\ t^*_{\alpha,\beta}(B,C)\le t^*_{\alpha,\beta}(A,C)$$

for every $A\subseteq B\subseteq C\in IFS(X)$. Then, the corresponding similarity measure is introduced by

$$s^*_{\alpha,\beta}(A,B)\stackrel{\text{def}}{=}1-t^*_{\alpha,\beta}(A,B)$$

Since the universe X is finite, we may also choose a suitable weight $\omega(\alpha, \beta) \in (0, +\infty)$, which depends on the mentioned before two arbitrary vectors: n-dimensional vector α (with components lying in $(0, 1]$) and k-dimensional vector β (with components lying in $(0, 1]$), such that for two arbitrary IFSs A and B, defined over the universe X, it is fulfilled:

$$\omega(\alpha, \beta) t_{\alpha,\beta}(A, B) \in [0, 1].$$

Then the mapping given by

$$t^{**}_{\alpha,\beta} \stackrel{\text{def}}{=} \omega(\alpha, \beta) t_{\alpha,\beta}$$

is also a distance – normalized distance, i.e.

$$t^{**}_{\alpha,\beta} : IFS(X) \times IFS(X) \to [0, 1].$$

Then, we construct the similarity measure, corresponding to the normalized distance $t^{**}_{\alpha,\beta}$:

$$s^{**}_{\alpha,\beta}(A, B) \stackrel{\text{def}}{=} 1 - t^{**}_{\alpha,\beta}(A, B)$$

Of course, there are also other ways for introducing similarity measures over the universe X, corresponding to appropriate distances in the same universe. Such possibilities will be a topic of a future investigation of ours.

Acknowledgements

This paper is partially supported by Grant BIn-2/09 "Design and development of intuitionistic fuzzy logic tools in information technologies" of the Bulgarian National Science Fund.

References

[1] Atanassov, K. Intuitionistic Fuzzy Sets, Springer Physica-Verlag, Heidelberg, 1999.

[2] Wang, X. Distance measure between intuitionistic fuzzy sets. Pattern Recognition Letters 26 (2005) 2063–2069.

[3] Szmidt, E., J. Kacprzyk. Distances between intuitionistic fuzzy sets. Fuzzy Sets Systems Vol. 114, 2000, No. 3, 505–518.

[4] Ye, J. Cosine similarity measures for intuitionistic fuzzy sets and their applications. Mathematical and Computer Modelling, Vol. 53, 2011, 91–97.