

New measures of similarity based on fuzzy implications

Erick González-Caballero^a, Susana Díaz^b Rafael Espín^c and Susana Montes^{b,*}

^aHigher Polytechnic Institute of Havana, Cuba

E-mail: erickgc@yandex.com

^bDepartment of Statistics and Operational Research, University of Oviedo, Spain

E-mails: {diazsusana,montes}@uniovi.es

^cAutonomous University of Coahuila, Mexico

E-mail: rafaelalejandroespinandrade@gmail.com

Abstract. An important problem in fuzzy theory is how to determine if two elements are similar or not. There are different definitions to measure how similar or close are elements. One of the most important concepts in this context is that one of similarity. This paper aims at studying fuzzy similarities defined by fuzzy implications, logical equivalences expressed by fuzzy bi-implication operators and aggregation operators. We propose a definition of proximity between two fuzzy finite sets by using the previous operators and we study when this type of operators satisfy the formal definition of similarity.

Keywords: Similarity measure, Fuzzy set, Logical equivalence, Fuzzy implication.

1. Introduction

The procedure of determining how close or similar (or the converse, how different) two elements appear in many fields like pattern recognition, clustering and classification (see, e.g., [29,30]). Thus, many proposals to overcome this problem can be found in the literature (see, e.g., [1,6,9,10,19,21,22,30]). A wide range of definitions were introduced to measure proximity (or the converse, differences) between samples. Among them we can cite distances, divergences or dissimilarities or their opposite, similarities, are one of those basic definitions when dealing with fuzzy sets. In this paper we have focused on the study of similarity relations obtained from implications.

One option in the fuzzy approach to similarities is the measure based on the geometric distance model (see [9]), where two vectors of fuzzy truth values are compared by means of some metric measures, usually the Minkowski r -metric defined for $r \geq 1$, which in-

cludes well-known metrics like Hamming (or Manhattan), Euclidean and Tchebychev (or sup distance), but other metric measures can be included (see [17]).

Other ways to define fuzzy similarities using measures based on the set theoretic approach, where classical fuzzy intersection, union and complement are used or by measures based on the matching function S , where a vector function is defined (see [9]).

The fuzzy generalization of the crisp equivalence relation can also be found, (see for example, [1] and [5]). There, the classical axioms of equivalence relations, reflexivity, symmetry and transitivity, are fuzzified.

In few words, there exists a great amount of similarity measures in the literature, defined from different criteria. However, no one of them proves to be adequate for all purposes [24]. For example, geometric measures like Minkowski metric are useful when *a priori* objective dimensions have to be measured, but they are not effective for other purposes like comparing faces, countries or personalities [31].

The main lack of geometric measures is they have to

*Corresponding author. E-mail: montes@uniovi.es.

satisfy a metric axiomatic, which is difficult to fulfill [17,27].

Feature-measure models discard the necessity that similarity measure be a metric. Nevertheless, they can not deal with intra-dimensional similarities. Besides, they admit representation of ordinal data, but they don't model the order [27].

Surveys about similarity measures, their advantages and drawbacks can be found in [27,31].

In this paper we follow the fuzzy logic approach to define similarity, which can be found in [17,18] by using fuzzy implications, the logical bi-implication and aggregation operators. Also, in [4], by defining the fuzzy operators from the concept of algebra of residuated lattice. In [2] we can find some notions of this approach applied to pattern recognition.

In this paper we propose measuring similarity between fuzzy sets using t-norms, t-conorms and negations together with fuzzy implications, logical equivalences expressed by fuzzy bi-implication operators and aggregation operators. We also study under which conditions that measure satisfies the definition of similarity.

In the literature, this kind of similarity measures are based on T-equalities [11,17]. A T-equality is a T-equivalence, and hence includes the axiom of T-transitivity. This axiom constrains similarity measures to be defined by R-implications.

This paper aims to demonstrate that the substitution of T-transitivity by monotonicity, like in [19], expands the scope of similarities based on bi-implications to an unexplored field. Nevertheless, we will show that not all bi-implications are similarity measures.

The advantages of our approach are the following:

- Fuzzy similarity measures based on bi-implications generalize the formulas of propositional logic. Therefore, the results maintain a logical perspective. This is an improvement on geometric measures.
- R-implications are defined with the supremum of a set, hence comparing for example with S-implications, they are less simple to calculate.
- It is uncommon to find a parameterized family of R-implications, maybe because of its definition. We found an exception in [28]. However in this paper we shall obtain easily other parameterized bi-implications.

It is important for experimental adjusting in decision making and in Similarity-based reasoning [2,24].

- We shall prove that the axiomatic we use is not independent of this kind of measures. This property simplifies the way to obtain new similarity measures based on bi-implications.
- This approach allows us to model dimensional similarities in the objects, like with geometric measures, and also its features. That is intrinsic to fuzzy sets.

Once characterized those t-norms that allow our measure to satisfy the definition of similarity, we classify several families of t-norms into those that satisfy these axioms. The results obtained allow us to select t-norms and aggregation operators to obtain the best similarity measure in each case.

Finally, we illustrate with an image processing example the applicability of the measure introduced.

The structure of the paper is the following: in Section 2 the basic concepts we use in the paper are summarized. In Section 3 the axiomatic of similarity that we will use along the paper and our measure of similarity based on logical equivalence are presented. Moreover, we present some theoretical results about similarity when dealing with R-implications, S-implications, QL-implications and f and g-generated implications, respectively. In Section 4 one example of image processing is used to illustrate the applicability of our theory. Finally, we address some conclusions in Section 5.

2. Basic concepts

In this section we will recall some basic notions and properties that will be useful hereafter. We will consider a finite universe X and we will denote $F(X)$ the set of fuzzy subsets of X and $A, B, C \dots$ the fuzzy subsets of X .

2.1. Operators

Definition 1 ([7,8,14]) *An aggregation operator is a function $\text{Aggreg} : \bigcup_{n \in \mathbb{N}} [0, 1]^n \rightarrow [0, 1]$ that satisfies the following conditions:*

1. *Boundary conditions.* $\text{Aggreg}(0, 0, \dots, 0) = 0$ and $\text{Aggreg}(1, 1, \dots, 1) = 1$.
2. *Symmetry.* For $(a_1, \dots, a_i, \dots, a_j, \dots, a_n) \in [0, 1]^n$, $\text{Aggreg}(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = \text{Aggreg}(a_1, \dots, a_j, \dots, a_i, \dots, a_n)$.
3. *Monotonicity.* If $a_1 = b_1, \dots, a_i \leq b_i, \dots, a_n = b_n$, then $\text{Aggreg}(a_1, \dots, a_i, \dots, a_n) \leq \text{Aggreg}(b_1, \dots, b_i, \dots, b_n)$.

An operator is called compensatory if it satisfies the following axiom (see [12]): $\text{Aggreg}(a_1, \dots, a_n)$ belongs to the interval $[\min(a_1, \dots, a_n), \max(a_1, \dots, a_n)]$, for any $a_1, \dots, a_n \in [0, 1]$.

Definition 2 (see [12]) Let $f : [0, 1] \rightarrow [0, \infty]$ be a real valued, continuous, strictly increasing function. A Quasi-arithmetic mean is defined by the formula:

$$M(a_1, a_2, \dots, a_n) = f^{-1} \left[\sum_{i=1}^n \frac{1}{n} f(a_i) \right].$$

Hereafter, we will use the definition above for the compensatory operators appearing in this paper:

Example 1 A kind of Quasi-arithmetic mean family is obtained when we consider $f(a) = a^p$, that is,

$$M_p(a_1, a_2, \dots, a_n) = \left[\sum_{i=1}^n \frac{1}{n} a_i^p \right]^{1/p}, p \in \mathbb{R},$$

where we agree that for $p < 0$, $\infty^p = 0$ and $0^p = \infty$. Let us note that if $p = 0$, M_p is not defined, but it tends to the geometric mean. Moreover, M_{-1} is the harmonic mean and M_1 is the arithmetic mean.

T-norms and t-conorms (see [3]) are particular cases of aggregation operators. They are usually considered in the context of fuzzy sets to formalize intersection and union respectively.

Definition 3 A t-norm is a binary operation $T : [0, 1]^2 \rightarrow [0, 1]$, such that for $a, b, c \in [0, 1]$ the following four axioms are satisfied:

1. $T(a, b) = T(b, a)$. (Commutativity)
2. $T(a, T(b, c)) = T(T(a, b), c)$. (Associativity)
3. $T(a, b) \leq T(a, c)$ whenever $b \leq c$. (Monotonicity)
4. $T(a, 1) = a$. (Boundary condition)

The most important examples of t-norms are: The minimum, $T_{\mathbf{M}}(a, b) = \min(a, b)$, the (algebraic) product, $T_{\mathbf{P}}(a, b) = a \cdot b$, the Łukasiewicz t-norm, $T_{\mathbf{L}}(a, b) = \max(a + b - 1, 0)$ and the drastic product,

$$T_{\mathbf{D}}(a, b) = \begin{cases} 0 & \text{if } \max(a, b) < 1, \\ \min(a, b) & \text{otherwise.} \end{cases}$$

The minimum is the greatest t-norm and the drastic product is the smallest one.

An automorphism φ is a strictly increasing continu-

ous function $\varphi : [0, 1] \rightarrow [0, 1]$ such that $\varphi(0) = 0$ and $\varphi(1) = 1$. Given a t-norm T and an automorphism φ , it holds that $T_{\varphi}(a, b) = \varphi^{-1}(T(\varphi(a), \varphi(b)))$ is also a t-norm. A t-norm is continuous if it is continuous as a function on its two components. Continuous t-norms are characterized as follows (see [16]): every continuous t-norm is either a φ -transformation of the Łukasiewicz t-norm or a φ -transformation of the product or an ordinal sum of t-norms in the two previous families. An ordinal sum of the countable family of t-norms $\{[a_s, b_s], T_s, s \in S\}$ is a new t-norm defined as follows:

$$T(a, b) = \begin{cases} a_s + (b_s - a_s) T_s \left(\frac{a - a_s}{b_s - a_s}, \frac{b - a_s}{b_s - a_s} \right) & \text{if } (a, b) \in [a_s, b_s]^2, \\ \min(a, b) & \text{otherwise.} \end{cases}$$

Definition 4 Let T be a t-norm.

- An element $a \in]0, 1[$ is called a nilpotent element of T if there exists some $n \in \mathbb{N}$ such that $a_T^{(n)} = 0$.
- An element $a \in]0, 1[$ is called a zero divisor of T if there exists some $b \in]0, 1[$ such that $T(a, b) = 0$.

It is well-known that each nilpotent element a of a t-norm T is also a zero divisor of T . If $n > 1$ is the smallest integer such that $a_T^{(n)} = 0$ then $T(a, a_T^{(n-1)}) = 0$ with $a_T^{(n-1)} > 0$, but not conversely.

Another family of operators very useful and related to t-norms is the family of t-conorms.

Definition 5 A t-conorm is a binary operation $S : [0, 1]^2 \rightarrow [0, 1]$, such that for $a, b, c \in [0, 1]$ the following four axioms are satisfied:

1. *Commutativity.* $S(a, b) = S(b, a)$.
2. *Associativity.* $S(a, S(b, c)) = S(S(a, b), c)$.
3. *Monotonicity.* $S(a, b) \leq S(a, c)$ whenever $b \leq c$.
4. *Boundary condition.* $S(a, 0) = a$.

T-conorms are dual operators of t-norms. For any t-norm T , the operator $S(a, b) = 1 - T(1 - a, 1 - b)$ is a t-conorm and the other way around, for any t-conorm, S the operator $T(a, b) = 1 - S(1 - a, 1 - b)$ is a t-norm. This way, the most important t-conorms are the maximum, $S_{\mathbf{M}}(a, b) = \max(a, b)$, the product t-conorm, $S_{\mathbf{P}}(a, b) = a + b - a \cdot b$, the Łukasiewicz t-conorm, $S_{\mathbf{L}}(a, b) = \min(a + b, 1)$ and the drastic product t-conorm,

$$S_{\mathbf{D}}(a, b) = \begin{cases} 1 & \text{if } \min(a, b) > 0, \\ \max(a, b) & \text{otherwise.} \end{cases}$$

The maximum is the smallest t-conorm and the drastic product t-conorm is the greatest one.

Definition 6 A negation is a decreasing function $N : [0, 1] \times [0, 1] \rightarrow [0, 1]$ such that $N(0) = 1$ and $N(1) = 0$. A negation is called strong if it is continuous, strictly decreasing and involutive, this is, if $N(N(a)) = a$.

The standard negation, defined as $N(a) = 1 - a$ for all $a \in [0, 1]$, is a strong negation.

2.2. Fuzzy implications

Fuzzy implications appeared as the fuzzy counterpart of the classical implication: $p \rightarrow q$. Since there are different ways of expressing this condition in the classical case, many different proposals appeared for the fuzzy counterpart in the literature as can be seen in [3]. One of the most commonly accepted axiomatic definition is the following one:

Definition 7 ([13,20]) A map $I : [0, 1] \times [0, 1] \rightarrow [0, 1]$ is called a fuzzy implication if it satisfies the following conditions:

- $I(\cdot, b)$ is decreasing for any fixed $b \in [0, 1]$,
- $I(a, \cdot)$ is increasing for any fixed $a \in [0, 1]$ and
- $I(0, 0) = I(1, 1) = 1$ and $I(1, 0) = 0$.

We next recall some of the most important families of fuzzy implications, starting by R -implications that are related to the formalism of Boolean logic (see [3, 13] among many others).

Definition 8 Let T be a t-norm, the implication defined by

$$I_T(a, b) = \sup\{c \in [0, 1] \mid T(a, c) \leq b\}$$

is called an R -implication, (see [3]).

Implications in the classical logic can also be written as $\neg p \vee q$ (not p or q). The following family tries to generalize this expression to the fuzzy context.

Definition 9 An S -implication associated with a t-conorm S and a strong negation N is defined by $I_{S,N}(a, b) = S(N(a), b)$, $\forall a, b \in [0, 1]$.

We will consider the most usual negation, the standard negation, $N(a) = 1 - a$, $\forall a \in [0, 1]$ and therefore the notation will be simplified to I_S .

Yet another way to express implications in the classical or crisp context is $\neg p \vee (p \wedge q)$ (not p or both p and q) and this is the departing point for defining QL -implications:

Definition 10 A QL -implication associated with a t-norm T , a t-conorm S and a fuzzy negation N is a function $I : [0, 1]^2 \rightarrow [0, 1]$ defined by $I_{T,S}(a, b) = S(N(a), T(a, b))$, $a, b \in [0, 1]$.

Along the paper, when we work with QL -implications, we will consider for each t-norm its dual t-conorm.

Definition 11 (see [3]) Let $f : [0, 1] \rightarrow [0, \infty]$ be a strictly decreasing and continuous function with $f(1) = 0$. The function $I_f : [0, 1]^2 \rightarrow [0, 1]$ defined by $I_f(a, b) = f^{-1}(af(b))$, $a, b \in [0, 1]$, with the understanding $0 \cdot \infty = 0$, is called an f -generated implication. The function f itself is called an f -generator of I_f .

Definition 12 (see [3]) Let $g : [0, 1] \rightarrow [0, \infty]$ be a strictly increasing and continuous function with $g(0) = 0$. The function $I_g : [0, 1]^2 \rightarrow [0, 1]$ defined by $I_g(a, b) = g^{(-1)}(\frac{1}{a}g(b))$, $a, b \in [0, 1]$, with the understanding $\frac{1}{0} = \infty$, $0 \cdot \infty = \infty$, where $g^{(-1)}(a)$ is the pseudo-inverse of g , given by:

$$g^{(-1)}(a) = \begin{cases} g^{-1}(a), & \text{if } a \in [0, g(1)], \\ 1, & \text{if } a \in [g(1), \infty]. \end{cases}$$

is called a g -generated implication. The function g itself is called a g -generator of I_g .

3. Measures of similarity based on fuzzy implications

For real numbers, if a is smaller than or equal to b and b is smaller than or equal to a , we can conclude that both numbers are equal. In fact, $a = b$ if and only if $a \leq b \wedge b \leq a$.

In a similar way, the connectives used in fuzzy logic, implications, can also be symmetrized to obtain a logical equivalence (see [25]):

$$E(a, b) = T(I(a, b), I(b, a)). \quad (1)$$

Taking into account this measure of equality between two numbers, a measure of the similarity or equality between two fuzzy sets was defined (see [2]) as follows:

$$E(A, B) = \text{Aggreg}_{x \in X}(E(A(x), B(x))). \quad (2)$$

Let us note that we will use the same formula for measuring similarity between two fuzzy subsets and for

the degree of equality between two values. There is no possible confusion since one of them refers to values in the interval $[0, 1]$ and the other one to fuzzy subsets of a universe X .

Definition 13 ([19]) Let $F(X)$ be the set of fuzzy subsets of X . A map $E : F(X) \times F(X) \rightarrow [0, 1]$ is a similarity measure if it satisfies, for any $A, B, C \in F(X)$, the following axiomatic:

1. Reflexivity. $E(A, A) = 1$.
2. Symmetry. $E(A, B) = E(B, A)$.
3. Monotonicity. $E(A, B) \geq E(A, C)$ and $E(B, C) \geq E(A, C)$ if $A \subseteq B \subseteq C$.

We are interested in the properties of the operator defined in Eq. (2). In particular, we would like to check in which cases it is a similarity in the sense of Def. 13, since that definition is one of the most usual measures of comparison between fuzzy sets.

Obviously, the properties satisfied by Eq. (2) will depend on the operators involved, i.e., on the implication and the t-norm considered. We present a detailed study for different types of implications in the following sections. Previously, we will prove some general results.

Proposition 1 Eq. (2) satisfies reflexivity if and only if the associated implication operator satisfies $I(a, a) = 1$ for all $a \in [0, 1]$.

Proof Assume that I satisfies that $I(a, a) = 1$ for all $a \in [0, 1]$. Then for any subset $A \subseteq X$ and any element $x \in X$, it holds that $I(A(x), A(x)) = 1$. Therefore, $E(A, A) = \text{Aggreg}_{x \in X}(E(A(x), A(x)))$ and then $E(A, A) = \text{Aggreg}_{x \in X}(T(1, 1)) = 1$.

Conversely, if there exists a value $a \in [0, 1]$ such that $I(a, a) < 1$, we can consider the subset A of X defined by $A(x) = a$ for all $x \in X$. Then

$$\begin{aligned} E(A, A) &= \text{Aggreg}_{x \in X}(E(A(x), A(x))) \\ &= \text{Aggreg}_{x \in X}(T(I(a, a), I(a, a))) \\ &\leq \text{Aggreg}_{x \in X}(I(a, a)) = I(a, a) < 1. \end{aligned}$$

And this means that E is not reflexive.

It follows from here that the first property that an implication has to satisfy to allow E to be a similarity measure is that any value implies itself with degree 1. But we can say more, if the implication satisfies that property, then the operator E also satisfies the third axiom of similarities.

Proposition 2 Let E be an operator defined as in Eq. (2) and such that the associated implication I verifies $I(a, a) = 1$ for all $a \in [0, 1]$. Then E satisfies the third axiom of similarities.

Proof Observe that for any pair of values $a, b \in [0, 1]$ such that $a \leq b$, since I is increasing in the second argument, it holds that $I(a, b) \geq I(a, a) = 1$ whereas $I(a, b) = 1$. This implies $E(a, b) = T(I(a, b), I(b, a)) = I(b, a)$ for all $a \leq b$.

Let us now consider three subsets $A \subseteq B \subseteq C$ of a universe X . Take some $x \in X$ and denote $a = A(x)$, $b = B(x)$ and $c = C(x)$. Then it holds that $a \leq b \leq c$ with $a, b, c \in [0, 1]$. As we have proven above, in this case it holds that $E(a, b) = I(b, a)$, $E(b, c) = I(c, b)$ and $E(a, c) = I(c, a)$. On the other hand, since I is decreasing in the first component and increasing in the second one, then $E(a, b) = I(b, a) \geq I(c, a) = E(a, c)$ and $E(b, c) = I(c, b) \geq I(c, a) = E(a, c)$. Equivalently,

$$\begin{aligned} E(A(x), B(x)) &= I(B(x), A(x)) \geq I(C(x), A(x)) \\ &= E(A(x), C(x)) \\ E(B(x), C(x)) &= I(C(x), B(x)) \geq I(C(x), A(x)) \\ &= E(A(x), C(x)). \end{aligned}$$

These inequalities hold for any $x \in X$. Thus, the result follows from the monotonicity of aggregation operators.

The condition $I(a, a) = 1$ for any $a \in [0, 1]$ is sufficient to assure that the operator E satisfies the third condition of similarity, as we have just seen. However, it is not a necessary condition as we can see in the following example.

Example 2 The S -implicator obtained from the maximum t-conorm and the standard negation provides a counterexample. Let's suppose $a, b \in [0, 1]$ and $a \leq b$. Then $I_{S_M}(a, b) = S_M(1 - a, b)$ and $I_{S_M}(b, a) = S_M(1 - b, a)$. Since $1 - a \geq 1 - b$ we have that $I_{S_M}(a, b) \geq I_{S_M}(b, a)$ and therefore

$$E_{T_M}(a, b) = T_M(I_{S_M}(a, b), I_{S_M}(b, a)) = I_{S_M}(b, a).$$

Thus, if $a \leq b \leq c$ we have that

$$E_{T_M}(a, b) = I_{S_M}(b, a) \geq I_{S_M}(c, a) = E_{T_M}(a, c)$$

and

$$E_{T_M}(b, c) = I_{S_M}(c, b) \geq I_{S_M}(c, a) = E_{T_M}(a, c).$$

However, $I_{S_M}(a, a) = S_M(1 - a, a) = \max(1 - a, a) \neq 1$ for all $a \in (0, 1)$.

Proposition 3 Let E be an operator defined as in Eq. (2). Then E satisfies the second axiom of similarities.

Proof The proof is immediate by the commutativity fulfilled by any t-norm. Thus $E(A(x), B(x)) = E(B(x), A(x))$, $\forall x \in X$ and $\forall A, B \in F(X)$.

Theorem 1 Let E be an operator defined as in Eq. (2). It is a similarity if and only if the associated implication operator fulfills $I(a, a) = 1$ for all $a \in [0, 1]$.

Proof The proof is an immediate consequence of Propositions 1, 2 and 3.

3.1. R-implications

First of all, we are going to prove that when we consider an R-implication, the operator in Eq. (1) does not depend on the t-norm considered. Then, neither does the operator defined in Eq. (2).

Lemma 1 Let T be a t-norm and I_T the R-implication obtained from T . For any t-norm T_1 and for any pair of values $a, b \in [0, 1]$ it holds that

$$T_1(I_T(a, b), I_T(b, a)) = \min(I_T(a, b), I_T(b, a)).$$

Proof Observe that for any pair of values a, b we have that $a \leq b$ or $b \leq a$. Then $I_T(a, b) = 1$ or $I_T(b, a) = 1$. Assume without loss of generality that $I_T(a, b) = 1$. Then for any t-norm T_1 it holds that $T_1(I_T(a, b), I_T(b, a)) = T_1(1, I_T(b, a)) = I_T(b, a) = \min(I_T(a, b), I_T(b, a))$.

Apart from this property of stability, we can prove that if we consider an R-implication, the operator E is a similarity measure.

Proposition 4 Let T be a t-norm and I_T the R-implication defined from T . Then the operator $E : F(X) \times F(X) \rightarrow [0, 1]$ defined as

$$E_R(A, B) = \text{Aggreg}_{x \in X}(E_R(A(x), B(x)))$$

is a similarity relation, where E_R is the operator defined by

$$E_R(a, b) = T_1(I_T(a, b), I_T(b, a))$$

for any $a, b \in [0, 1]$ and for any t-norm T_1 .

Proof This follows from Theorem 1, since it is immediate that $I_T(a, a) = 1$ for all $a \in [0, 1]$.

Thus, we always obtain a similarity from an R-implication.

3.2. S-implications

If we replace R-implications by S-implications, we can also obtain a measure of equivalence (see [25]). The expression looks as follows in this case. For any pair of values $a, b \in [0, 1]$,

$$E_S(a, b) = T(I_S(a, b), I_S(b, a)),$$

where S is a t-conorm, I_S is the implication defined from S as in Def. 9 and the t-norm T is the dual t-norm of the t-conorm S .

This expression can be used in Eq. (2) to provide a new measure of the equivalence or proximity between two fuzzy sets.

$$E_S(A, B) = \text{Aggreg}_{x \in X} E_S(A(x), B(x)),$$

where A and B are two fuzzy subsets of the universe X .

Despite what happens with R-implications, in the context of S-implications the t-norm chosen changes the measure of equivalence or similarity considered, as we can see in the following example.

Example 3 Consider the t-conorm S_P and the values $a = 0.6, b = 0.5$ it holds that $I_{S_P}(a, b) = (1 - a) + b - (1 - a)b = (1 - 0.6) + 0.5 - (1 - 0.6) \cdot 0.5 = 0.7$ and $I_{S_P}(b, a) = (1 - b) + a - (1 - b)a = (1 - 0.5) + 0.6 - (1 - 0.5) \cdot 0.6 = 0.8$. Therefore,

$$\begin{aligned} E_{S_P}(0.6, 0.5) &= T(I_{S_P}(0.6, 0.5), I_{S_P}(0.5, 0.6)) \\ &= T(0.7, 0.8), \end{aligned}$$

and the final result depends on the t-norm chosen.

Moreover, now we cannot assure that $I_S(a, a) = 1$ for any $a \in [0, 1]$. In fact, there are t-conorms such that this is not true for any a .

Example 4 If we continue with the previous example, $I_{S_P}(a, a) = (1 - a) + a - (1 - a)a = 1 - a + a^2 \neq 1, \forall a \in (0, 1)$.

Thus, we cannot use Propositions 1 and 2. However, reflexivity is totally characterized for S-implications as follows.

Proposition 5 Let T be a t-norm. The associated logical equivalence E_S is reflexive if and only if $T(a, 1 - a) = 0$, for all $a \in [0, 1]$.

Proof We have proven in Theorem 1 that the operator E_S is reflexive if and only if $I_S(a, a) = 1$ for all $a \in [0, 1]$ and this happens if and only if $S(a, 1 - a) = 1$ for all $a \in [0, 1]$, equivalently, if and only if $T(a, 1 - a) = 0$ for all $a \in [0, 1]$.

It follows from the previous Proposition that only a particular type of t-norms with zero divisors allow us to construct similarities by means of S -implications. In particular, the previous result rules out the minimum and product t-norms. On the other side, Łukasiewicz and the nilpotent-minimum t-norms guarantee that the operator E is a similarity.

Remark 1 Thus, if the logical equivalence is reflexive, we have that the t-norm is such that the set of nilpotent elements is $]0, 1[$. Unfortunately, the converse is not true. Thus, for instance, the Schweizer-Sklar t-norm $T_{SS}(a, b) = \left(\max(\sqrt{a} + \sqrt{b} - 1, 0)\right)^2$ fulfills that any $a \in]0, 1[$ is a nilpotent element, since $T_{SS}(a, (1 - \sqrt{a})^2) = 0$ and $(1 - \sqrt{a})^2 \in [0, 1]$ for any $a \in]0, 1[$. However, $E_{T_{SS}}(0.3, 0.3) = 0.7163 \neq 1$.

The previous characterization allows us to identify which of the operators in the main families of t-norms (see [16]) allow us to build operators E_S that are similarities. We next classify them into operators that lead to similarities and those that do not.

If we first consider the four most important t-norms, the minimum, product, Łukasiewicz and the drastic t-norm, it is easy to verify that only Łukasiewicz and the drastic t-norms satisfy the condition $T(a, 1 - a) = 0$. Another t-norm that satisfies this condition is the nilpotent minimum defined as follows

$$T_{nM}(a, b) = \begin{cases} 0, & \text{if } a + b \leq 1 \\ \min(a, b), & \text{otherwise.} \end{cases}$$

Let us consider some of the most important families of t-norms.

The Yager Family of t-norms (see, [16]) is defined BY:

$$T_Y^\lambda(a, b) = \begin{cases} T_D(a, b), & \text{if } \lambda = 0 \\ T_M(a, b), & \text{if } \lambda = \infty \\ \max\left\{1 - \left[(1-x)^\lambda + (1-y)^\lambda\right]^{1/\lambda}, 0\right\}, & \text{if } \lambda \in (0, \infty) \end{cases}$$

The Sugeno-Weber t-norms are defined by:

$$T_{SW}^\lambda(a, b) = \begin{cases} T_D(a, b), & \text{if } \lambda = -1 \\ T_P(a, b), & \text{if } \lambda = \infty \\ \max\left(\frac{a+b-1+\lambda ab}{1+\lambda}, 0\right), & \text{if } \lambda \in (-1, \infty) \end{cases}$$

The Schweizer-Sklar family is defined by:

$$T_{SS}^\lambda(a, b) = \begin{cases} T_M(a, b), & \text{if } \lambda = -\infty \\ T_P(a, b), & \text{if } \lambda = 0 \\ T_D(a, b), & \text{if } \lambda = \infty \\ \left(\max(a^\lambda + b^\lambda - 1, 0)\right)^{1/\lambda}, & \text{if } \lambda \in \mathbb{R} - \{0\} \end{cases}$$

We have not recalled the complete definition of other families, since in those cases only one member of the family satisfies the condition $T(a, 1 - a) = 0$. An exhaustive list of these families and their properties can be found in [16]. In the family of Hamacher, only the drastic t-norm ($\lambda = 0$) satisfies $T(a, 1 - a) = 0$ and therefore only this operator allows E_S to be a similarity. The same happens if we consider the Mayor-Torrens family or the Frank family: only the Łukasiewicz t-norm, corresponding to the parameter $\lambda = 1$ and $\lambda = \infty$ respectively, satisfies the condition. In the case of Aczél-Alsina and Dombi's family, is the drastic t-norm, corresponding to $\lambda = 0$ in both cases, the one that satisfies the condition.

Next we discuss for the Yager, Sugeno-Weber and Schweizer-Sklar families which t-norms satisfy the axioms of similarity, equivalently, which of them satisfy $T(a, 1 - a) = 0$.

– We first consider the family of Yager. Condition $T(a, 1 - a) = 0$ is equivalent to solve the inequality $1 - \left[(1-a)^\lambda + a^\lambda\right]^{1/\lambda} \leq 0$ or $(1-a)^\lambda + a^\lambda \geq 1$. This inequality holds if and only if $\lambda \leq 1$. For any $\lambda > 1$ it holds that $0.5^\lambda < 0.5$ and therefore $0.5^\lambda + (1 - 0.5)^\lambda < 1$. On the other hand, for any $\lambda \leq 1$ it holds that $a^\lambda \geq a$ for all $a \leq 1$ and therefore, $a^\lambda + (1 - a)^\lambda \geq 1$ for all $a \in [0, 1]$.

– For the Sugeno-Weber family, the condition is $\frac{a+(1-a)-1+\lambda a(1-a)}{1+\lambda} \leq 0$, which is equivalent to $\lambda a(1-a) \leq 0$. Since $a(1-a) \geq 0$, the inequality only holds if $\lambda \leq 0$, this is, if $\lambda \in]-1, 0]$.

- For the Schweizer-Sklar family we have to find the values where the condition $a^\lambda + (1-a)^\lambda - 1 \leq 0$ or equivalently $a^\lambda + (1-a)^\lambda \leq 1$ is satisfied. It is easy to prove that this inequality is true for $\lambda \in [1, \infty[$.

Table 1 summarizes the families of t-norms which allow us to obtain similarities from S-implications, where S is the dual t-conorm associated with this t-norm. The interval where the definition of similarity is valid is also included.

The set of nilpotent t-norms which satisfy the axioms of similarity is not restricted to the t-norms included in Table 1, as we can see in the following example.

Families of t-norms	Generation of similarity
Schweizer-Sklar $\lambda \in [-\infty, \infty]$	$\lambda \in [1, \infty]$
Hamacher $\lambda \in [0, \infty]$	$\lambda = \infty$
Frank $\lambda \in [0, \infty]$	$\lambda = \infty$
Yager $\lambda \in [0, \infty]$	$\lambda \in [0, 1]$
Aczél-Alsina $\lambda \in [0, \infty]$	$\lambda = 0$
Dombi $\lambda \in [0, \infty]$	$\lambda = 0$
Sugeno-Weber $\lambda \in [-1, \infty]$	$\lambda \in [-1, 0]$
Mayor-Torrens $\lambda \in [0, 1]$	$\lambda = 1$
Nilpotent Minimum t-norm	Yes

Table 1

Families of t-norms and the nilpotent minimum t-norm (left-continuous). Parameters for which they satisfy the axiomatic of similarity.

Example 5 If we define $\varphi(a) = \frac{\tan(a)}{\tan(1)}$, φ is a strictly increasing and continuous function in the interval $[0, 1]$. Also $\varphi(0) = 0$ and $\varphi(1) = 1$. Therefore, the t-norm $T_{\tan}(a, b) = \varphi^{-1}(T_L(\varphi(a), \varphi(b)))$ is a nilpotent t-norm.

Now, since $\tan(x+y) = \frac{\tan(x)+\tan(y)}{1-\tan(x)\tan(y)} \geq \tan(x) + \tan(y)$ for any $x, y \in [0, 1]$, in particular $\tan(1) \geq \tan(x) + \tan(1-x)$ and therefore, $\varphi(1) \geq \varphi(x) + \varphi(1-x)$. Then the condition in Proposition 5 is fulfilled. Hence, from T_{\tan} we can define a measure of similarity. However, T_{\tan} does not belong to any of the families considered in Table 1.

3.3. QL-implications

The operator presented in Def. 10 is not necessarily an implication according to Def. 7, because $I(\cdot, b)$

is not always decreasing. One example can be found in [3]. They also provide a necessary condition for the t-conorm and negation involved in the definition in order for a QL-implication to satisfy all the axioms in Def. 7, in particular the monotonicity in its first argument:

Proposition 6 (see [3]) Let T be a t-norm, S a t-conorm, N a strong negation and I the operator given in Def. 10. If $I(\cdot, b)$ is decreasing then

$$S(x, N(x)) = 1 \text{ for all } x \in [0, 1].$$

Also in [3] we can find one example that proves that the condition in Proposition 6 is only necessary, but not sufficient.

However, our interest focuses on the study of similarities. We showed in Proposition 1 that in order for an operator E defined from an implication to be a similarity, the implication operator has to satisfy $I(a, a) = 1$ for all $a \in [0, 1]$.

The following proposition shows that this condition is not satisfied by any t-norm when applied to QL-implications.

Proposition 7 There exists no t-norm T such that the binary operation $E_{T,S}$ associated to the QL-implication $I_{T,S}$, where S is the dual t-conorm of T , is a similarity.

Proof Notice that $I_{T,S}(a, a) = 1$ is equivalent to $S(1-a, T(a, a)) = 1$ and this is equivalent to $T(a, 1-T(a, a)) = 0$ for all $a \in [0, 1]$. As a consequence, for every $a \leq 0.5$ it holds that $a + T(a, a) \leq a + a \leq 1$. Therefore, $a \leq 1 - T(a, a)$.

This implies, since t-norms are increasing, that $0 = T(a, 1 - T(a, a)) \geq T(a, a)$. Equivalently, $T(a, a) = 0$ for all $a \leq 0.5$. If we now apply the hypothesis to a value $0 < a < 0.5$ we get

$$1 = S(1-a, T(a, a)) = S(1-a, 0) = 1-a.$$

A contradiction.

The conclusion is that the set of t-norms which can be used for defining similarities according to Eq. (1) is empty in the case of QL-operators.

3.4. f and g-generated implications

We next study if f and g -generated implications lead to a similarity in the sense of Def. 13.

As we proved in Theorem 1, the implication operator I_f must satisfy the Identity Principle ($I(a, a) = 1$). However, there is no implication operator I_f generated from a function f that satisfies this equality except for $a = 0$ or $a = 1$ (see [3]). Therefore E_f , the measure of proximity defined as in Eq. (2), does not satisfy reflexivity and consequently, it is not a similarity.

Also in [3], we can find a characterization for those implication operators of the type I_g that satisfy the Identity Principle.

Proposition 8 [3] *The implication operator I_g satisfies $I_g(a, a) = 1$ for all $a \in [0, 1]$ if and only if $g(1) < \infty$ and $a \leq \frac{g(a)}{g(1)}$ for all $a \in [0, 1]$.*

As a direct corollary we can give the following result.

Corollary 1 *E_g , the measure of proximity defined as in Eq. (2) using a g -generated implication operator I_g is a similarity if and only if $g(1) < \infty$ and $a \leq \frac{g(a)}{g(1)}$ for all $a \in [0, 1]$.*

Example 6 *One example is the Yager's class of additive generators, $g^\lambda(a) = a^\lambda$ (see [3]), where $\lambda \in]0, \infty[$. $g^\lambda(1) = 1 < \infty$ for every λ and the pseudo-inverse is $(g^\lambda)^{(-1)}(a) = \min(1, a^{1/\lambda})$. It satisfies the inequality $a \leq \frac{g^\lambda(a)}{g^\lambda(1)}$ if $\lambda \in]0, 1]$. Hence, the implication operator is the following:*

$$I_{g^\lambda}(a, b) = \min\left(1, \frac{b}{a^{1/\lambda}}\right). \quad (3)$$

4. Application to image processing problems

Here, we illustrate the application of the measures of similarity to compare one original image obtained by contrast imaging in nuclear magnetic resonance with other three digital images processed from it. This example can be seen in [23].

A black and white image is digitally represented by a matrix, where each element is an integer value of the grayscale range 0-255 (see [26]). The a_{ij} element means the normalized gray value of the pixel placed in this position, 0 means the pixel is black and 1 means the pixel is white, intermediate values refer to a gray color. The nearer it is to 1, the lighter it is, and the opposite: the nearer it is to 0, the darker it is.

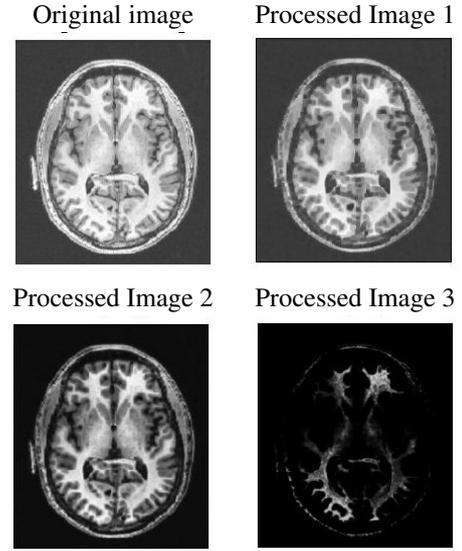


Fig. 1. The original and three processed images of one brain contrast imaging in nuclear magnetic resonance.

In this example the original figure (on the left) will be compared with the other three images, which were obtained by processing the first one. We will answer the question: How much does the original image look like the other three? To answer this question we will compare every pixel by the grayscale level. We will use two implications, one is the Gödel implication, an R -implication (see [3]) associated to the minimum t-norm and defined by:

$$I_G(a, b) = \begin{cases} 1, & \text{if } a \leq b \\ b, & \text{if } a > b \end{cases}$$

The other one we have chosen is the bi-implication defined by a S -implication from the Schweizer-Sklar t-norm with $\lambda = 2$. Finally, we used the g -generated implication from Yager's class for $\lambda = 1$ and T_M as the t-norm in Eq. (1).

We have compared images pairwise by using these two operators as implications. As aggregation operator we have tested three options: the t-norm T_M , the Geometric Mean and the Arithmetic Mean, as compensatory operators.

The results are summarized in the table below. Let us call A_0, A_1, A_2 and A_3 the matrices corresponding to the original, the first, the second and the third images, respectively.

The similarity between the original image and the other ones is considered 0 for the minimum and the

Equiv. def.	R-implication (Gödel)		
	T_M	GM	AM
$E(A_0, A_0)$	1	1	1
$E(A_0, A_1)$	0	0	0.51
$E(A_0, A_2)$	0	0	0.36
$E(A_0, A_3)$	0	0	0.17

Equiv. def.	S-implication (Schweizer-Sklar) $\lambda = 2$		
	T_M	GM	AM
$E(A_0, A_0)$	1	1	1
$E(A_0, A_1)$	0	0	0.96
$E(A_0, A_2)$	0	0	0.93
$E(A_0, A_3)$	0	0	0.61

Equiv. def.	g-generated (Yager's class) $\lambda = 1$		
	T_M	GM	AM
$E(A_0, A_0)$	1	1	1
$E(A_0, A_1)$	0	0	0.84
$E(A_0, A_2)$	0	0	0.50
$E(A_0, A_3)$	0	0	0.17

Table 2

Pairwise comparison of images, for the Gödel, the S-implication defined from the Schweizer-Sklar t-norm for $\lambda = 2$ and the Yager's class for $\lambda = 1$ implications, where the aggregation operators are T_M , the Geometric Mean and the Arithmetic Mean.

Geometric Mean due to the fact that some pixels are white in one image and black in the other one. This fact influences the final result of the operator evaluating the similarity between both images when some aggregation operators are considered. We can see that the arithmetic mean is not as affected as the other two operators and the similarity measure is higher in that case.

If we focus on the results obtained for the arithmetic mean, we can realize that they maintain an order that agrees with intuition. Those values confirm the visual impression, that the similarity between the original and the three others decreases from the first one to the third one (from left to right in Fig. 1).

Despite there is a coincidence in the order of classification according to every measure, the cardinality of the similarity measures are completely different, mainly the calculus corresponding to the arithmetic mean. The advantage of the parameterized similarity measures (the two on the right) is that their values can be adjusted to the criterion of a decision maker. This is a

benefit to be remarked. Furthermore, they are obtained in a simple way and the computational cost of calculation is very low. These advantages become essentials if we could design a Similarity Based Reasoning.

For example, in the similarity measure defined from the S-implication with the arithmetic mean in Table 2, for $\lambda = 1$ we obtained $E(A_0, A_1) = 0.8967$ and for $\lambda = 3$, $E(A_0, A_1) = 0.9700$, and we could select the value of λ which best fits the decision maker's criteria. Also, it is clear that if we choose a continuous implication, its corresponding similarity measure will preserve the intra-dimensional similarities and at the same time they can be represented according to linguistic variables, for example 'lightness' or 'darkness'.

5. Conclusions

In this paper one definition of similarity based on logical bi-implication equivalence is studied. We use a classic axiomatic for similarity between two fuzzy sets to demonstrate that our definition of similarity is valid. From the results obtained, we can do the following conclusions:

- The equivalence definition based on an R-implication is a similarity measure for any t-norm.
- The equivalence definition based on an S-implication has to use a t-norm satisfying the condition $T(a, 1 - a) = 0$. The classical t-norms which satisfy this condition are: the drastic t-norm, Łukasiewicz, the nilpotent minimum t-norm, the Schweizer-Sklar t-norm family for $\lambda \in [1, \infty]$, the Sugeno-Weber family for $\lambda \in [-1, 0]$ and the Yager family for $\lambda \in [0, 1]$. Since Łukasiewicz and the drastic t-norms are particular members of the following families, we can also say that the Frank family and the Hamacher family satisfy the condition for $\lambda = \infty$, the Mayor-Torrens family for $\lambda = 1$ and Aczel-Alsina and Dombi's families for $\lambda = 0$.
- QL-implications do not generate any similarity measure, according to the definitions we used in this paper. Reflexivity is the condition that fails.
- Bi-implications defined from f-generated implications do not satisfy the axiom of reflexivity. Besides, g-generated implications can be used to define a measure of similarity if and only if $g(1) < \infty$ and $\forall a \ a \leq \frac{g(a)}{g(1)}$.
- The theory developed in this paper can be applied to compare digital images. We have illustrated this with one example.

- Our approach has many advantages: A family of parameterized bi-implications with low computational cost can be obtained. Also, the measures can be used in quantitative similarity calculus with qualitative meanings.

Acknowledgements

The research reported in this paper has been partially supported by the Spanish Ministerio de Economía y Competitividad under Project TIN2014-59543-P and EUREKA SD project (agreement number 2013-2591), that is supported by the Erasmus Mundus programme of the European Union.

References

- [1] S. Ashraf. Fuzzy dissimilarity and generalization of Valverde's theorem on T-indistinguishability relations. *Fuzzy Sets and Systems* 275, 144–154, 2015.
- [2] L. Baccour, W. Maghrebi, S. Kanoun and A. M. Alimi, A comparative study of a fuzzy similarity measure aggregated with fuzzy implications applied to shape recognition. In: *proceedings of the 6th Mexican International Conference on Artificial Intelligence, Proc. of MICAI*, 2007.
- [3] M. Bacczyński and B. Jayaram. *Fuzzy Implications*. Springer, 2008.
- [4] R. Belohlavek and V. Vychodil. Similarity issues in attribute implications from data with fuzzy attributes. In: *In Information Reuse and Integration, 2006 IEEE International Conference*. pp. 132–135, 2006.
- [5] U. Bodenhofer, *A similarity-based generalization of fuzzy orderings*. Vol. 26, Linz, Austria: Universitätsverlag Rudolf Trauner, 1999.
- [6] B. Bouchon-Meunier, M. Rifqi and S. Bothorel. Towards general measures of comparison of objects, *Fuzzy Sets and Systems* 84, 143–153, 1996.
- [7] H. Bustince, J. Fernandez, A. Kolesárová and R. Mesiar. Generation of linear orders for intervals by means of aggregation functions. *Fuzzy Sets and Systems* 220, 69–77, 2013.
- [8] T. Calvo, G. Mayor, and R. Mesiar, *Aggregation operators: new trends and applications*. Vol. 97, Physica, 2012.
- [9] S. M. Chen, M. S. Yeh and P. Y. Hsiao. A comparison of similarity measures of fuzzy values. *Fuzzy Sets and Systems* 72, 79–89, 1995.
- [10] I. Couso, L. Garrido and L. Sánchez. Similarity and dissimilarity measures between fuzzy sets: A formal relational study. *Information Sciences*, 229, 122–141, 2013.
- [11] B. De Baets and R. Mesiar. Metrics and T-Equalities, *Journal of Mathematical Analysis and Applications*, 267, 531–547, 2002.
- [12] M. Detyniecky, Doctoral Thesis at University Paris VI, *Mathematical Aggregation Operators and their Application to Video Querying*, 2000.
- [13] J. Fodor and M. Roubens. *Fuzzy preference modelling and multicriteria decision support*. Kluwer Academic Publishers, 1994.
- [14] M. Grabisch, J. L. Marichal, R. Mesiar and E. Pap. Aggregation functions: Means. *Information Sciences* 181, 1–22, 2011.
- [15] P. Hájec, *Mathematics of Fuzzy Logic*, Kluwer, 1998.
- [16] E. P. Klement, R. Mesiar and E. Pap. *Triangular Norms*. Kluwer Academic Publishers, 2000.
- [17] H. Le Capitaine. A relevance-based learning model of fuzzy similarity measures. *IEEE Transactions on Fuzzy Systems* 20, 57–68, 2012.
- [18] H. Le Capitaine and C. Frelicot. Towards a unified logical framework of fuzzy implications to compare fuzzy sets. In: *In International Fuzzy Systems Association World Congress and European Society for Fuzzy Logic and Technologies Conference*. pp. 1200–1205, 2009.
- [19] X. Lui. Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets and Systems* 52, 305–318, 1992.
- [20] S. Massanet and J. Torrens. The law of importation versus the exchange principle on fuzzy implications. *Fuzzy Sets and Systems* 168, 47–69, 2011.
- [21] I. Montes, N. R. Pal, V. Janiš and S. Montes. Divergence Measures for Intuitionistic Fuzzy Sets, *IEEE Transactions on Fuzzy Systems* 232, 444–456, 2015.
- [22] S. Montes, I. Couso, P. Gil and C. Bertoluzza. Divergence measure between fuzzy sets, *International Journal of Approximate Reasoning* 30, 91–105, 2002.
- [23] P. Quirós, P. Alonso, I. Díaz, A. Jurío, S. Montes. An hybrid construction method based on weight functions to obtain interval-valued fuzzy relations. *Mathematical Methods in the Applied Sciences*, DOI: 10.1002/mma.3443, 2015.
- [24] S. Raha, A. Hossain, S. Ghosh. Similarity based approximate reasoning: fuzzy control. *Journal of Applied Logic* 6, 47–71, 2008.
- [25] J. Recasens. *Indistinguishability Operator*. Springer, 2010.
- [26] S. J. Sangwine and R.E. Horne. *The colour image processing handbook*. Springer Science & Business Media, 2012.
- [27] A. Schwering. Approaches to Semantic Similarity Measurement for Geo-Spatial Data: A Survey. *Transactions in GIS* 12, 5–29, 2008.
- [28] T. Whalen. Parameterized R-implications. *Fuzzy sets and systems*, 134, 231–281, 2003.
- [29] N. Xiong. Learning fuzzy rules for similarity assessment in case-based reasoning. *Expert Systems with Applications* 38, 10780–10786, 2011.
- [30] J. Ye. Cosine similarity measures for intuitionistic fuzzy sets and their applications. *Mathematical and Computer Modelling* 53, 91–97, 2011.
- [31] R. Zwick, E. Carstein and D. V. Budescu. Measures of Similarity Among Fuzzy Concepts: A Comparative Analysis. *International Journal of Approximate Reasoning* 1, 221–242, 1987.