Acta Universitatis Matthiae Belii, series Mathematics
Volume 30 (2022), 36–77, ISSN 1338-7111
Available at https://actamath.savbb.sk

Tools for comparison of fuzzy sets

Vladimír Kobza
Department of Mathematics, Faculty of Natural Sciences, Matej Bel University
Tajovského 40, 974 01 Banská Bystrica, Slovak Republic
vladimir.kobza@umb.sk

Abstract

The basic study of fuzzy sets theory was introduced by Lotfi Zadeh in 1965. Many authors investigated possibilities how two fuzzy sets can be compared and the most common kind of measures used in the mathematical literature are dissimilarity measures. The previous approach to the dissimilarities is too restrictive, while there exist many pairs of fuzzy sets, which are incomparable to each other with respect to inclusion $\subseteq$. Therefore we need some new concept for measuring a difference between fuzzy sets so that it could be applied for arbitrary fuzzy sets. In order to overcome this problem the divergence measures as another kind of measures useful to compare two fuzzy sets have been introduced recently. The last condition in previous dissimilarity measures $\max\{D(A,B),D(B,C)\} \leq D(A,B)$ has been replaced by the new one $\max\{D(A\cup C,B\cup C),D(A\cap C,B\cap C)\} \leq D(A,B)$, where the fuzzy sets operations - union and intersection are defined by a triangular conorm S and its dual triangular norm T, respectively. In this paper we focus on the special class of so-called local divergences and their possible generalizations.

Received: December 28, 2022
Accepted: January 20, 2023
Available online: January 26, 2023
Communicated with Miroslav Haviar.

Keywords fuzzy set, triangular norm (conorm), divergence measure, dissimilarity measure, distance measure, local divergence, generalized local divergence, entropy measure.

MSC(2020) Primary: 03B52 Fuzzy logic, logic of vagueness; 03E72 Theory of fuzzy sets; Secondary: 03E75 Applications of set theory.

1 Introduction

This survey gives an overview on the approaches to the measuring of differences between fuzzy sets. A comparison between two fuzzy sets can be done by different comparison measures, by considering different points of view. In some cases these are measures of the equality degree (e.g. similarity measures) and in other cases the difference degree (e.g. dissimilarity measures, divergence measures and distance measures). A basic study about measures of comparison was given by Bouchon-Meunier et al. in *Towards general measures of comparison of objects* (1996), see [3]. More recently, a review of the measures based on the differences, where the relationships among them were studied in detail, was proposed by Couso et al. in *Similarity and dissimilarity measures between fuzzy sets: A formal relational study* (2013), see [5]. Among these measures, divergences appear as a good alternative in some cases:

- Montes (1998): Partitions and divergence measures in fuzzy models [12],
- Montes et al. (2002): Divergence measure between fuzzy sets [15],
- Kobza et al. (2017): Generalizated local divergence measures [7].

The main aims of the study published in [7] were the possible generalizations of local divergences in three directions (Section 11):

(1) the fuzzy sets operations (union and intersection) are defined by means of arbitrary triangular norm (conorm), i.e. we consider the divergences associated to the triple (X, T, S),

(2) for each element of the universal set some specific importance is assigned, i.e. we consider the different maps h_x for each element $x \in X$ in definition of local divergence,

(3) we have extended the family of local divergences into the richer family of generalized local divergences based on some triangular conorm S. We say that the divergence measure D is S-local if for arbitrary $A, B \in \mathcal{F}(X)$ and crisp set $Z \in \mathcal{P}(X)$ we have $D(A, B) = S[D(A \cup Z, B \cup Z), D(A \cup Z^c, B \cup Z^c)]$.

During many years of the mentioned research on divergence measures not only a number of applications have been created, but also other related more general theoretical models, mostly based on real life situations:

- *intuitionistic fuzzy sets (IF-sets):* combining the information on the membership and non-membership of an object to a collection. Formally, an IF-set could be understood as a pair of ordinary fuzzy sets, but as the membership and non-membership functions are closely related, they are studied as a single object [11].
- *hesitant fuzzy sets:* this approach is given by the following motivation. A group of evaluators decides about the membership of an object to a collection, the result of their decision is an n-tuple of their opinions. In case they work together and exchange information on their decision process, these decisions cannot be considered as independent and thus they should be considered as a single object. Hence, multisets (see [2]) are proper objects for the values of such decision processes. Technically, if we do not care about the order of the evaluators, we do not need to work with n-tuples (where the order is important). On the other hand, it would be limiting to work with subsets (of all possible decisions), as the number of evaluators with identical decisions is usually important. This idea leads directly to the notion of a hesitant fuzzy set [8].

Pattern recognition and decision making are some of the areas where fuzzy sets find their important applications. Naturally, estimation of similarity between sets (patterns or alternatives) is crucial in many aspects. This task is usually performed by applying a suitable comparison measure to the pair of fuzzy sets corresponding to given patterns or alternatives. The aim was to suggest some possible applications of the previous theoretical results for the practical areas of interests as pattern recognition and decision making (Section 12).

For definitions and basic facts used throughout this paper we refer the reader to the monographs [4, 6].

2 Fuzzy sets

The basic study related to fuzzy logic and fuzzy sets theory has been introduced by Lotfi Aliasker Zadeh in 1965 ([22, 23]). Zadeh proposed mathematical tools for modeling a special type of an uncertainty related to an object with fuzzy boundaries. However, some contributions to the many valued logic appeared earlier, for example by the Polish mathematician Jan Łukasiewicz. Zadeh's ideas have been accepted not only because the development of the computer sciences reached an important level but it was also influenced by development of the social and technological progress in general.

Let X denote the **universal set** (or shortly only the universe). The family of all subsets of X, or equivalently the power set, will be denoted by $\mathcal{P}(X)$. For each element $x \in X$ and each subset $A \subseteq X$ only two cases are possible: $x \in A$ or $x \notin A$. Therefore, the characteristic function, denoted by χ_A, is introduced as follows:

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A; \\ 0 & \text{if } x \notin A. \end{cases}$$

In the fuzzy sets theory we will require that each element $x \in X$ belongs to the set A with some degree from the unit interval $[0, 1]$, i.e. not only the degree 0 for $x \notin A$ and 1 for $x \in A$. We give an example. Let X denote the set of all flights on the world (X is finite set in this case) and the subset A be formulated as the set of all long-distance flights. So, the flight between Madrid and Santiago de Chile belongs to the set A in a relatively high degree. However, also the flight between Madrid and Vienna can be considered as a long-distance since other local connections are taking into account, and therefore it also belongs to the set A, but with much lower degree.

It makes sense for each element $x \in X$ to assign the value (the **membership degree**), denoted by $\mu_A(x)$. Now we can define the **membership function** in the following way:

$$\mu_A : X \to [0, 1], x \mapsto \mu_A(x).$$

The pair (X, μ_A) is said to be a **fuzzy set**.

Obviously, the membership function μ_A is a generalization of the characteristic function χ_A. So, each set $A \in \mathcal{P}(X)$ is a special case of a fuzzy subset of X. Each fuzzy set A is uniquely determined by its membership function μ_A and vice versa. Therefore, the notions A and μ_A may be identified. Usually, the notion μ_A will be not used in our text and instead of $\mu_A(x)$ we will write $A(x)$ for a membership degree of an element $x \in X$ to the fuzzy set A.

The family of all fuzzy subsets on the universe X will be denoted by $\mathcal{F}(X)$. More formally,

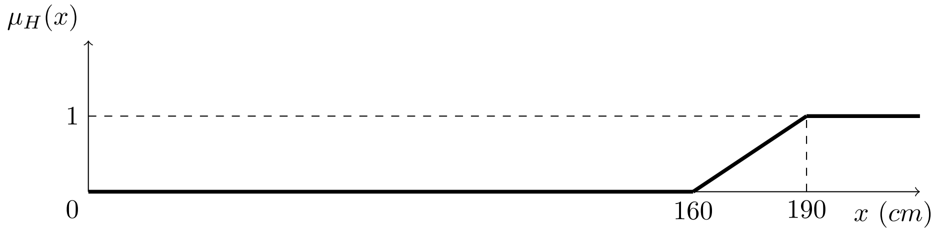
$$\mathcal{F}(X) := \{\mu_A \mid \mu_A : X \to [0, 1]\} = [0, 1]^X.$$

Evidently, $\mathcal{F}(X) \supseteq \mathcal{P}(X)$. Also the empty set $\varnothing$ and the universal set X can be considered as the fuzzy sets with membership functions defined as:

$$\mu_\varnothing(x) = 0 \text{ and } \mu_X(x) = 1 \text{ for all } x \in X.$$

We give one motivation for modelling of real situations by fuzzy sets theory. Let us consider the notion "high person". There is no exact boundary, which can be considered to decide if a person is high or not. Someone can say 160 cm, someone else 180 cm, etc. We can define a fuzzy set H on the universal set $X = (0, \infty)$ (of course, just in the theoretical way) by the following membership function μ_H:

$$\mu_H(x) = \begin{cases} 0, & \text{if } x \in (0, 160), \\ \frac{1}{30}(x - 160), & \text{if } x \in [160, 190], \\ 1, & \text{if } x \in (190, \infty). \end{cases}$$



In this case we have the following membership degrees: $\mu_H(160\,cm) = 0$, $\mu_H(175\,cm) = 0.5$, $\mu_H(190\ cm) = 1$. Of course, this modelling depends on the society, geographical conditions, regions and therefore, it is very subjective.

Next, we define some important notions related to fuzzy sets.

Definition 1. Let $A \in \mathcal{F}(X)$. Then:

- the **support** of the fuzzy set A, denoted by $Supp(A)$ is

$$Supp(A) = \{x \in X \mid \mu_A(x) > 0\}\,,$$

- the **kernel** of the fuzzy set A, denoted by $Ker(A)$ is

$$Ker(A) = \{x \in X \mid \mu_A(x) = 1\}\,,$$

- the **height** of the fuzzy set A, denoted by $hgt(A)$ is

$$hgt(A) = \sup_{x \in X} \mu_A(x),$$

- a fuzzy set A is said to be **normal** if there exists $x \in X$, such that $\mu_A(x) = 1$, or equivalently $Ker(A) \neq \varnothing$,
- a fuzzy set A is said to be **subnormal** if A is not normal.

The ordering on $\mathcal{F}(X)$ is defined in the following way:

Definition 2. Let $A, B \in \mathcal{F}(X)$. Then:

- $A \subseteq B \Leftrightarrow \mu_A(x) \leq \mu_B(x)$ for all $x \in X$,
- $A = B \Leftrightarrow A \subseteq B$ and $B \subseteq A$, i.e. $\mu_A(x) = \mu_B(x)$ for all $x \in X$.

For each element $x \in X$ the membership degree $A(x)$ has been assigned. It is said to be a **vertical representation** of the fuzzy set A. However, we can also assign for each $\alpha \in [0,1]$ some subset of X consisting of all elements having their membership degrees greater or equal to α. We define the α**-cut**, denoted by A_α, as follows:

$$A_\alpha = \{x \in X \mid A(x) \geq \alpha\}\,.$$

Obviously, we obtain $A_0 = X$ and $A_1 = Ker(A)$ as special cases for $\alpha \in \{0,1\}$. The family of all α-cuts will be denoted by $\{A_\alpha\}_{0 \leq \alpha \leq 1}$. Each fuzzy set can be decomposed into the family of α-cuts, i.e. elements of $\mathcal{P}(X)$. We call it a **horizontal representation** of the fuzzy set A.

Let $A \in \mathcal{F}(X)$. If we know the family of α-cuts, then for each $x \in X$ the membership degree $A(x)$ as the supremum of all $\alpha \in [0,1]$ such that $x \in A_\alpha$ can be determined. More formally, using the characteristic function for A_α it can be written as follows:

$$A(x) = \sup_{0 \leq \alpha \leq 1} \min\{\alpha, \chi_{A_\alpha}(x)\}\,.$$

3 Triangular norms and triangular conorms

Triangular norms have been introduced into the mathematical literature by Karl Menger in 1942. Triangular norms and conorms are operations which generalize the conjunction and disjunction in fuzzy logic. They were originally used to generalize the triangle inequality from classical metric spaces to probabilistic metric spaces. The principal studies can be found in [17], [18] or [19] by Berthold Schweizer and Abe Sklar. In the original axioms for triangular norms no associativity was required. Theory of continuous t-norms has two rather independent roots, namely, the field of functional equations and the theory of topological semigroups (see [6]). The development of t-norms was impacted by János Aczél's monograph [1] published in 1966. The full characterization of continuous Archimedean t-norms by means of additive generators has been done after 1960 by Ling (see [9]) and Schweizer and Sklar (see [17]). Some parametric families of t-norms, firstly Frank's family, have been studied later.

Triangular norms will be mentioned in this section. These functions are useful for modeling a conjunction in fuzzy logic and intersection of fuzzy sets.

Definition 3. The **triangular norm** (t-norm) is a function $T : [0,1] \times [0,1] \to [0,1]$ satisfying the following conditions:

(T1) $T(a,b) = T(b,a)$, for all $a,b \in [0,1]$ (commutativity),

(T2) $T(T(a,b),c) = T(a,T(b,c))$, for all $a,b,c \in [0,1]$ (associativity),

(T3) $b \leq c \Rightarrow T(a,b) \leq T(a,c)$, for all $a,b,c \in [0,1]$ (monotonicity),

(T4) $T(a,1) = a$, for all $a \in [0,1]$ (boundary condition).

Therefore, the function T is a monotone, associative and commutative operation defined on $[0,1] \times [0,1]$ with neutral element 1. Some important examples of t-norms, so-called basic t-norms, are the following:

- Minimum t-norm: $T_M(a,b) = \min(a,b)$, for all $a,b \in [0,1]$,
- Product t-norm: $T_P(a,b) = a \cdot b$, for all $a,b \in [0,1]$,
- Łukasiewicz t-norm: $T_L(a,b) = \max(a+b-1,0)$, for all $a,b \in [0,1]$,
- Drastic t-norm:

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

For these basic t-norms, it holds that $T_D \leq T_L \leq T_P \leq T_M$. In fact, for any t-norm T, it is fulfilled that $T_D \leq T \leq T_M$.

Moreover, for any t-norm T the following properties are fulfilled:

- $T(a,0) = 0$, for all $a \in [0,1]$,
- $T(a,b) \leq \min\{a,b\}$, for all $a,b \in [0,1]$.

To show these relations we notice:

- from the monotonicity condition (T3) and the boundary condition (T4) from Definition 3 we have that: $0 \leq T(a,0) \leq T(1,0) = 0$, for all $a \in [0,1]$, and therefore $T(a,0) = T(0,a) = 0$,

- using again the conditions (T3) and (T4) from Definition 3 we have that: $T(a,b) \le T(a,1) = a$ and $T(a,b) \le T(1,b) = b$, simultaneously, for all $a, b \in [0,1]$.

Definition 4. Let T_1, T_2 be t-norms. If for all pairs $(a,b) \in [0,1]^2$,

$$T_1(a,b) \le T_2(a,b),$$

we say that t-norm T_1 is **weaker** than the t-norm T_2, or equivalently, T_2 is **stronger** than T_1.

Changing the neutral element from 1 to 0, we obtain the triangular conorm (t-conorm), a function used for modeling the disjunction in fuzzy logic and union of fuzzy sets.

Definition 5. The **triangular conorm** (t-conorm) is a function $S : [0,1] \times [0,1] \to [0,1]$ satisfying the following conditions:

(S1) $S(a,b) = S(b,a)$, for all $a, b \in [0,1]$ (commutativity),

(S2) $S(S(a,b),c) = S(a,S(b,c))$, for all $a,b,c \in [0,1]$ (associativity),

(S3) $b \le c \Rightarrow S(a,b) \le S(a,c)$, for all $a,b,c \in [0,1]$ (monotonicity),

(S4) $S(a,0) = a$, for all $a \in [0,1]$ (boundary condition).

Thus, a t-conorm has all properties of triangular norms as monotonicity, associativity and commutativity, but it differs in the neutral element, which is now 0.

Similarly, for any t-conorm S the following properties are fulfilled:

- $S(a,1) = 1$, for all $a \in [0,1]$,
- $S(a,b) \ge \max\{a,b\}$, for all $a,b \in [0,1]$.

The t-norm T and t-conorm S are said to be dual if for each $a,b \in [0,1]$ the equation $T(a,b) = 1 - S(1-a, 1-b)$ is fulfilled.

For each previous example of basic t-norm we can consider its dual basic t-conorm as follows:

- Maximum t-conorm: $S_M(a,b) = \max(a,b)$, for all $a,b \in [0,1]$,
- Probabilistic sum: $S_P(a,b) = a + b - a \cdot b$, for all $a,b \in [0,1]$,
- Łukasiewicz t-conorm: $S_L(a,b) = \min(a+b, 1)$, for all $a,b \in [0,1]$,
- Drastic t-conorm:

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

For the basic t-conorms it holds: $S_D \ge S_L \ge S_P \ge S_M$. For any t-conorm S it holds: $S_D \ge S \ge S_M$.

Next, we introduce some important notions related to triangular norms (conorms).

Definition 6. Let T be a t-norm. If $T(a,a) = a$ for some element $a \in [0,1]$, the element a is said to be an **idempotent** element of T.

Obviously $T(0,0) = 0$ and $T(1,1) = 1$ for each t-norm T, i.e. the elements $0, 1$ are both idempotent; they are named the **trivial** idempotent elements. However, t-norms T_P and T_L have no other idempotent elements, although for T_M each element from unit interval $[0,1]$ is idempotent:

$$T_L(a,a) = a \Leftrightarrow \max\{2a-1, 0\} = a \Leftrightarrow a = 0 \textit{ or } a = 1,$$

$$T_P(a,a) = a \Leftrightarrow a^2 = a \Leftrightarrow a = 0 \textit{ or } a = 1,$$

$$T_M(a,a) = a \Leftrightarrow \min\{a,a\} = a \textit{ for all } a \in [0,1].$$

We have defined a triangular norm (conorm) as the map $T(S) : [0,1]^2 \to [0,1]$ fulfilling four conditions. It is possible to extend these maps as follows: one can define $T : [0,1]^n \to [0,1]$ (and analogously S) as

$$T(a_1, \ldots, a_n) = T(T(a_1, \ldots, a_{n-1}), a_n).$$

In special case for $n = 1$ we define $T(a) = a$. It allows us to introduce the concept of generalized local divergences, which are aggregated by arbitrary t-conorm.

For the basic t-norms we have:

$$T_M(a_1, \ldots, a_n) = \min\{a_1, \ldots, a_n\},$$

$$T_P(a_1, \ldots, a_n) = \prod_{i=1}^{n} a_i,$$

$$T_L(a_1, \ldots, a_n) = \max\left\{\sum_{i=1}^{n} a_i - (n-1), 0\right\},$$

$$T_D(a_1, \ldots, a_n) = \begin{cases} a_i, & \text{if } a_j = 1 \text{ for all } j \neq i, \\ 0, & \text{otherwise.} \end{cases}$$

Analogously we can extend the basic t-conorms S dual to T, i.e. $S(a_1, \ldots, a_n) = 1 - T(1-a_1, \ldots, 1-a_n)$:

$$S_M(a_1, \ldots, a_n) = \max\{a_1, \ldots, a_n\},$$

$$S_P(a_1, \ldots, a_n) = 1 - \prod_{i=1}^{n}(1 - a_i),$$

$$S_L(a_1, \ldots, a_n) = \min\left\{\sum_{i=1}^{n} a_i, 1\right\},$$

$$S_D(a_1, \ldots, a_n) = \begin{cases} a_i, & \text{if } a_j = 0 \text{ for all } j \neq i, \\ 1, & \text{otherwise.} \end{cases}$$

At the end of this section some remarks related to generating of triangular norms (conorms) will be given. We start with the definition of continuity and archimedean property.

Definition 7. Let T be a triangular norm and $(a_0, b_0) \in [0,1]^2$. We say that the map T is **continuous at a point** (a_0, b_0) if $\lim_{(a,b)\to(a_0,b_0)} T(a,b) = T(a_0, b_0)$. Moreover, T is **continuous** in $[0,1]^2$ if it is continuous at each point $(a_0, b_0) \in [0,1]$.

Obviously, the functions T_M, T_P, T_L are all continuous, but T_D is not. For example, for the sequences $a_n = \frac{n}{n+1}, b_n = \frac{1}{2}$ we have a_n approaches 1 and b_n approaches $\frac{1}{2}$ for $n \to \infty$. Thus, we have:

$$\lim_{n\to\infty} T_D(a_n, b_n) = \lim_{n\to\infty} T_D\left(\frac{n}{n+1}, \frac{1}{2}\right) = 0 \neq \frac{1}{2} = T_D\left(1, \frac{1}{2}\right) = T_D\left(\lim_{n\to\infty} a_n, \lim_{n\to\infty} b_n\right),$$

and therefore T_D is not continuous.

Definition 8. The triangular norm T (or conorm S) is said to be **archimedean** if for each $a, b \in [0,1]$ there exists $n \in \mathbb{N}$ such that $a_T^n < b$ (or $a_S^n > b$), where $a_T^n = T(a, \dots, a), a_S^n = S(a, \dots, a)$ and $(a, \dots, a) \in [0,1]^n$.

For example, T_P, T_L, T_D are all archimedean, but T_M is not, since we have $T_M(a,a) = a$ for each $a \in [0,1]$, but $a < a$ is not fulfilled.

One important property of some triangular norms will be remarked. The t-norms, which are continuous and archimedean, can be generated by a function $f : [0,1] \to \mathbb{R}_0^+$ satisfying the following conditions:

(i) f is decreasing,

(ii) f is continuous,

(iii) $f(1) = 0$.

Two cases $f(0) = c \in \mathbb{R}$ or $f(0) = \infty$ can be considered. We can construct an inverse map $f^{-1} : [0,c] \to [0,1]$, which is also decreasing. Since $c < \infty$, the domain of the map f^{-1} can be extended into $\mathbb{R}_0^+$, and therefore the map $f^{(-1)} : \mathbb{R}_0^+ \to [0,1]$ can be defined in the following way:

$$f^{(-1)}(a) = \begin{cases} f^{-1}(a), & \text{if } a \leq c, \\ 0, & \text{if } a > c. \end{cases}$$

The map $f^{(-1)}$ is said to be an **pseudo-inverse** to the function f. Obviously, if $c = \infty$, then $f^{(-1)} = f^{-1}$.

The following proposition is crucial. The proof can be found in [6] and [19].

Proposition 9. *The function $T : [0,1] \to [0,1]$ is a continuous archimedean t-norm if and only if there exists a continuous decreasing function $f : [0,1] \to \mathbb{R}_0^+$ with boundary condition $f(1) = 0$, such that for each pair $(a,b) \in [0,1]^2$ the following holds:*

$$T(a,b) = f^{(-1)}(f(a) + f(b)),$$

where $f^{(-1)}$ is a pseudo-inverse to the function f.

The function f is said to be an **additive generator** of the t-norm T.

The continuous archimedean triangular conorms can be generated analogously. We consider the function $g : [0,1] \to \mathbb{R}_0^+$, for which the following conditions are satisfied:

(i) g is increasing,

(ii) g is continuous,

(iii) $g(0) = 0$.

Also in this case we define the map $g^{(-1)} : \mathbb{R}_0^+ \to [0,1]$ as follows:

$$g^{(-1)}(a) = \begin{cases} g^{-1}(a), & \text{if } a \leq f(1), \\ 1, & \text{if } a > f(1). \end{cases}$$

Similarly, each continuous archimedean t-conorm can be generated by function g fulfilling the previous conditions such that:

$$S(a,b) = g^{(-1)}(g(a) + g(b)).$$

Finally, we remark the following notation. Triangular norm (conorm) is said to be:

- **nilpotent**, if $f(0) \in \mathbb{R}$ $(g(1) \in \mathbb{R})$,
- **strict**, if $f(0) = \infty$ $(g(1) = \infty)$.

4 Fuzzy set operations

Another important concepts will be the containment relation and the complement set. In particular we will consider the standard Zadeh's negation for defining the complement (see [22]).

Definition 10. Let $A, B \in \mathcal{F}(X)$. The **complement** of A denoted by A^c is the fuzzy set $A^c(x) = 1 - A(x)$, for all $x \in X$, and A is said to **be contained** in B, which is denoted by $A \subseteq B$, if $A(x) \leq B(x)$ for all $x \in X$.

Apart from the previous relation of containment, we need to consider the concepts of intersection and union of fuzzy sets. The initial definitions were also given by Zadeh (see [22]) as follows:

- intersection of A and B: $(A \cap B)(x) = \min(A(x), B(x))$, for all $x \in X$.
- union of A and B: $(A \cup B)(x) = \max(A(x), B(x))$, for all $x \in X$.

These are the standard operations, since they were considered at the initial definition. We can notice that they coincide with the usual operations for crisp sets. However, they are not the only way to generalize the corresponding classical set operations, but there exists a broad class of functions to represent them. For the intersection, this class consists of t-norms and for the union it consists of t-conorms. The remaining part of this section will be devoted to these concepts.

Using t-norms and t-conorms, we can define in general the intersection and union of two fuzzy as follows.

Definition 11. Let $A, B \in \mathcal{F}(X)$. Given a t-norm T and a t-conorm S,

- the intersection of A and B with respect to T is defined as the fuzzy set whose membership function is $A \cap_T B(x) = T(A(x), B(x))$, for all $x \in X$;
- the union of A and B with respect to S is defined as the fuzzy set whose membership function is $A \cup_S B(x) = S(A(x), B(x))$, for all $x \in X$.

Thus, we can denote by (X, T, S) the triple formed by the universe with the t-norm and the t-conorm defining the intersection and the union, respectively.

Usually we write shortly $\cup, \cap$ instead of $\cup_S, \cap_T$ if it is clear what triple (X, T, S) is considered.

Not only union or intersection are sufficient for dealing with fuzzy set operations. A way to create a complement A^c of the fuzzy set A will be discussed in this section. We start with the following definition of a fuzzy negation.

Definition 12. The **fuzzy negation** is a function $n : [0,1] \to [0,1]$ satisfying the following conditions:

(n1) $n(0) = 1\ ,\ n(1) = 0$ (boundary conditions),

(n2) $a < b \Rightarrow n(a) \geq n(b)$, for all $a, b \in [0,1]$ (monotonicity).

As we can see, each fuzzy negation is decreasing. In special case some stronger conditions for fuzzy negation can be considered:

Definition 13. Len n be the fuzzy negation satisfying the conditions (n1) and (n2). The decreasing and continuous fuzzy negation n is called a **strict fuzzy negation**. The involutive fuzzy negation n, i.e. such that $n(n(a)) = a$ for all $a \in [0,1]$, is called a **strong fuzzy negation**.

These negations are related. Each strong negation is strict, the converse implication does not hold.

Proposition 14. *Let n be the strong fuzzy negation. Then n is strict.*

Proof. To show that the fuzzy negation n is strict, i.e. decreasing and continuous, it suffices to prove that n is bijective map.

(1) We show that n is injective. Let $a, b \in [0,1]$ such that $n(a) = n(b)$. Then $n(n(a)) = n(n(b))$. Since n is involutive, we have $a = b$, i.e. n is injective. Since n is injective and decreasing, it is strictly decreasing.

(2) We show that n is surjective, i.e. for all $b \in [0,1]$ there exists $a \in [0,1]$ such that $b = n(a)$. Let us suppose $a = n(b)$. Then $n(a) = n(n(b)) = b$, since n is involutive. Since n is surjective and strictly decreasing, we have that n is continuous. Suppose that n is not continuous at a point $c \in [0,1]$. Then $\lim_{x\to c^-} n(x) \neq \lim_{x\to c^+} n(x)$. Since n is decreasing by (1), then $\lim_{x\to c^-} n(x) > \lim_{x\to c^+} n(x)$. It means that any point from the open interval

$$\left(\lim_{x\to c^-} n(x), \lim_{x\to c^+} n(x) \right)$$

has no inverse image, which contradicts the fact that n is surjective.

□

Now, we introduce two basic examples of a strong fuzzy negation.

Example 15. The fuzzy negation $n(a) = 1 - a$ is a strong fuzzy negation. The function n obviously fulfills the boundary (n1) and the monotonicity (n2) conditions. Moreover, for all $a \in [0,1]$ we have $n(n(a)) = 1 - (1 - a) = a$, i.e. the fuzzy negation n is strong.

Example 16. The fuzzy negation $n(a) = \sqrt{1-a^2}$ is a strong fuzzy negation. Also in this case, the function n fulfills the boundary (n1) and the monotonicity (n2) conditions. Since

$$n(n(a)) = \sqrt{1 - \sqrt{1-a^2}^2} = \sqrt{1-(1-a^2)} = \sqrt{a^2} = a$$

for all $a \in [0,1]$, we conclude that the fuzzy negation n is strong.

Definition 17. Let $A \in \mathcal{F}(X)$ and n be a strong fuzzy negation. Then the fuzzy set A^c defined as

$$A^c(x) = n(A(x)),$$

for all $x \in X$, is said to be a **fuzzy complement** to the fuzzy set A.

Once we have introduced the basic operations between fuzzy sets, we are able to start to work on the comparison of fuzzy sets.

5 Dissimilarity measures

Definition 18. A map $D : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ is a **dissimilarity measure** if it satisfies the following axioms:

(1) $D(A,A) = 0$ for every $A \in \mathcal{F}(X)$;

(2) $D(A,B) = D(B,A)$ for every $A, B \in \mathcal{F}(X)$;

(3) for every $A, B, C \in \mathcal{F}(X)$ such that $A \subseteq B \subseteq C$, it holds that $D(A,C) \geq \max(D(A,B), D(B,C))$.

There are several examples of dissimilarities. As this definition is not too restrictive, it is possible to define a counterintuitive measure of comparison for which the above axioms hold. The restriction associated to this definition is given by the fact that the requirement in Axiom (Diss.3) is only given for sets such that $A \subseteq B \subseteq C$, but there are a lot of sets which are not comparable with respect to $\subseteq$ and therefore, nothing is required for them. Thus, we need a concept where the restrictions about "proximity" are given for any sets.

6 Divergence measures

In order to overcome this problem, another measure of comparison between fuzzy sets was proposed in [12], the divergence measure, which satisfies the following natural properties:

- it becomes zero when the two sets coincide,
- it is a non-negative and symmetric function,
- it decreases when the two subsets become "more similar" in some sense.

While it is easy to formulate the first and the second conditions analytically, the third one depends on the formalization of the concept of "more similar". We base our approach on the fact that if we add (in the sense of union) a subset C to both fuzzy subsets A, B, we obtain two subsets which are closer to each other; the same for the intersection. So we propose the following:

Definition 19. Let (X, T, S) be a triple with X a universe and T and S any t-norm and t-conorm, respectively. A map $D : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ is a **divergence measure** with respect to (X, T, S) if for all $A, B \in \mathcal{F}(X)$, D satisfies the following conditions:

(1) $D(A, A) = 0$;
(2) $D(A, B) = D(B, A)$;
(3) $\max\{D(A \cup C, B \cup C), D(A \cap C, B \cap C)\} \leq D(A, B)$, for all $C \in \mathcal{F}(X)$, where the union and intersection are defined by means of S and T, respectively.

It is clear that a divergence measure is associated to a triple (X, T, S) and a map D can be a divergence measure with respect to one t-norm and need not be a divergence measure with respect to another t-norm. However, when there is not ambiguity, we will say just a divergence measure without specifying the used t-norm and t-conorm. We present some examples of divergence measures.

Example 20.
$$D(A, B) = \begin{cases} 0, & \text{if } A = B. \\ 1, & \text{if } A \neq B. \end{cases}$$
It is easy to see that D is a divergence. The first and the second conditions are trivial.

If $A \neq B$, then $D(A, B) = 1$ and $D(A \cap C, B \cap C) \in \{0, 1\} \leq D(A, B)$. If $A = B$, then $D(A, B) = 0$ and $A \cap C = B \cap C$ and by definition $D(A \cap C, B \cap C) = 0 \leq D(A, B)$. The same holds for $D(A \cup C, B \cup C)$. We can conclude that D is a divergence for any (X, T, S).

Example 21. We consider the triple (X, T_M, S_M), where X is a finite universe. For any pair of fuzzy sets in X we define the function D using the **Hamming distance** as follows:
$$D(A, B) = \sum_{x \in X} \alpha_x \cdot |A(x) - B(x)|,$$
where $\alpha_x \geq 0$ for any $x \in X$ and $\sum_{x \in X} \alpha_x = 1$.

Again, the first and second condition are trivial. Denote $A(x) = a, B(x) = b, C(x) = c$. If without loss of generality we assume $a \geq b$, then $T(a, c) \geq T(b, c)$ for any T, in particular for T_M. Thus,

- if $c > a \geq b$, then $|T_M(a, c) - T_M(b, c)| = |a - b| \leq |a - b|$,
- if $a \geq c \geq b$, then $|T_M(a, c) - T_M(b, c)| = |c - b| \leq |a - b|$,
- if $a \geq b > c$, then $|T_M(a, c) - T_M(b, c)| = |c - c| = 0 \leq |a - b|$,

and therefore, in all the cases, this inequality holds. From here we have that

$$\begin{aligned} D(A \cap C, B \cap C) &= \sum_{x \in X} \alpha_x \cdot |(A \cap C)(x) - (B \cap C)(x)| \\ &= \sum_{x \in X} \alpha_x \cdot |T_M(A(x), C(x)) - T_M(B(x), C(x))| \\ &\leq \sum_{x \in X} \alpha_x \cdot |A(x) - B(x)| = D(A, B). \end{aligned}$$

Analogously we prove that $D(A \cup C, B \cup C) \leq D(A, B)$ if we consider the maximum t-conorm for defining the union of two fuzzy sets. Thus, D is a divergence.

D is also a divergence if we consider the product t-norm or the Łukasiewicz t-norm and their dual t-conorms for defining the corresponding operations between sets, that is, if we work on (X, T_P, S_P) or (X, T_L, S_L). This is true since

- $|T_P(a,c) - T_P(b,c)| = |a \cdot c - b \cdot c| = c \cdot |a-b| \leq |a-b|$ and $|S_P(a,c) - S_P(b,c)| = |a + c - a \cdot c - (b + c - b \cdot c)| = (1-c) \cdot |a-b| \leq |a-b|$;
- if $a+c \geq 1$ and $b+c \geq 1$, then $|T_L(a,c) - T_L(b,c)| = |(a+c-1) - (b+c-1)| = |a-b|$ and $|S_L(a,c) - S_L(b,c)| = |1-1| = 0 \leq |a-b|$;
- if $a+c < 1$ and $b+c \geq 1$, then $|T_L(a,c) - T_L(b,c)| = |0 - (b+c-1)| = |b-(1-c)|$, but $a < 1-c \leq b$ and therefore $|T_L(a,c) - T_L(b,c)| \leq |b-a| = |a-b|$; moreover, $|S_L(a,c) - S_L(b,c)| = |a+c-1| = |(1-c)-a| \leq |b-a|$;
- the case $a+c \geq 1$ and $b+c < 1$ is analogously to the previous one;
- if $a+c < 1$ and $b+c < 1$, then $|T_L(a,c) - T_L(b,c)| = |0-0| = 0 \leq |a-b|$ and $|S_L(a,c) - S_L(b,c)| = |(a+c) - (b+c)| = |a-b|$.

However, this does not hold in general. For instance, if we consider the drastic t-norm, for the case $a = 1,\ b = 0.2,\ c = 0.9$ we have that $|T_D(a,c) - T_D(b,c)| = |T_D(1,0.9) - T_D(0.2,0.9)| = |0.9 - 0| = 0.9 > 0.8 = |a-b|$. Then, $D(A,B) = \sum_{x \in X} \alpha_x \cdot 0.8 = 0.8 > 0.9 = D(A \cap C, B \cap C)$.

It also follows from the definition that a divergence measure is always non-negative.

Proposition 22. *Let D be a divergence measure for any triple (X,T,S). It holds that $D(A,B) \geq 0$ for all $A,B \in \mathcal{F}(X)$.*

Proof. For any pair of elements A and B in $\mathcal{F}(X)$, we have that $D(A,B) \geq D(A \cap \varnothing, B \cap \varnothing) = D(\varnothing,\varnothing) = 0$. □

The following proposition emphasizes the fact that if two fuzzy sets become more closer (in sense of inclusion), then their divergence is smaller.

Proposition 23. *Let D be a divergence measure and A_1, A_2, B_1 and B_2 be the fuzzy sets such that $A_1 \subseteq B_1 \subseteq B_2 \subseteq A_2$. Then $D(B_1,B_2) \leq D(A_1,A_2)$.*

Proof. Since we work on (X, T_M, S_M), the following hold: $B_1 = B_1 \cap B_2 = A_1 \cup B_1, B_2 = A_2 \cap B_2, A_2 = A_2 \cup B_1$. Using the fact that D is a divergence measure and using the third property from Definiton 19, we can conclude: $D(B_1,B_2) = D(B_1 \cap B_2, A_2 \cap B_2) \leq D(B_1, A_2) = D(A_1 \cup B_1, A_2 \cup B_1) \leq D(A_1, A_2)$. □

By the last axiom of Definition 19, we could think that conditions $D(A \cap C, B \cap C) \leq D(A,B)$ and $D(A \cup C, B \cup C) \leq D(A,B)$ are equivalent. However, this is not true in general, even in the case T and S are dual, as we can see in the following counterexample.

Example 24. We will show that even if Axioms (Diss1) and (Diss2) in Definition 19 are fulfilled, the two conditions in Axiom (Div3) are not equivalent. We will start with an example where the condition for the union is fulfilled but it is not fulfilled for the intersection. Later another example in the converse direction will be shown. In both cases we will work on (X, T_M, S_M).

Let $D : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ be a function defined by:

$$D(A,B) = \sum_{x \in X} h(A(x), B(x)),$$

where

$$h(x,y) = \begin{cases} 0, & \text{if } x = y. \\ 1 - xy, & \text{otherwise.} \end{cases}$$

Axiom (Diss2) from the definition of divergence is satisfied since the function h is symmetric. By definition $h(x,x)=0$ and therefore Axiom (Diss1) is also satisfied.

We continue with Axiom (Div3). Let us consider the following partition of the universal set $X = X_1 \cup X_2 \cup X_3 \cup X_4 \cup X_5 \cup X_6$:

$$\begin{aligned} X_1 &= \{x \in X, A(x) \le B(x) \le C(x)\},\ X_2 = \{x \in X, A(x) \le C(x) < B(x)\},\\ X_3 &= \{x \in X, B(x) < A(x) \le C(x)\},\ X_4 = \{x \in X, B(x) \le C(x) < A(x)\},\\ X_5 &= \{x \in X, C(x) < A(x) \le B(x)\},\ X_6 = \{x \in X, C(x) < B(x) < A(x)\}.\end{aligned}$$

Then

$$D(A,B) = \sum_{i=1}^{6} \sum_{x\in X_i} h(A(x),B(x)).$$

Applying the standard maximum t-conorm we have:

$$D(A\cup C, B\cup C) = \sum_{x\in X_1\cup X_3} h(C(x),C(x)) + \sum_{x\in X_2} h(C(x),B(x))$$

$$+ \sum_{x\in X_4} h(A(x),C(x)) + \sum_{x\in X_5\cup X_6} h(A(x),B(x)).$$

Since h is, by definition, a positive and decreasing function in both components, we have that:

- $h(C(x),C(x)) = 0 \le h(A(x),B(x))$ for all $x \in X_1 \cup X_3$,
- $h(C(x),B(x)) < h(A(x),B(x))$ for all $x \in X_2$,
- $h(A(x),C(x)) < h(A(x),B(x))$ for all $x \in X_4$,
- $h(A(x),B(x)) = h(A(x),B(x))$ for all $x \in X_5 \cup X_6$.

Finally,

$$D(A\cup C, B\cup C) \le \sum_{x\in X} h(A(x),B(x)) = D(A,B).$$

We have shown that $D(A\cup C, B\cup C) \le D(A,B)$. However, the inequality for the intersection is not fulfilled in general. Thus, if we consider the universal set formed by only one element, i.e. $X=\{x\}$, the dual minimum t-norm T_M and the fuzzy sets A,B,C defined as $A(x)=0.2, B(x)=0.8, C(x)=0.5$, we have that $D(A\cap C, B\cap C) = h((A\cap C)(x),(B\cap C)(x)) = h(T_M(A(x),C(x)),T_M(B(x),C(x))) = h(A(x),C(x)) = 1-0.2\cdot 0.5 = 0.9$ and $D(A,B) = h(A(x),B(x)) = 1-0.2\cdot 0.8 = 0.84$, and therefore $D(A\cap C, B\cap C) > D(A,B)$.

In the previous example we have shown that $D(A\cup C, B\cup C) \le D(A,B)$ does not imply $D(A\cap C, B\cap C) \le D(A,B)$ in general. If we change the definition of D, then we can show that even the converse implication does not hold. We continue with an example where the condition for the intersection is fulfilled but it is not for the union.

Example 25. Let $D : \mathcal{F}(X)\times\mathcal{F}(X) \to \mathbb{R}$ be a function defined by

$$D(A,B) = \sum_{x\in X} h(A(x),B(x)),$$

where

$$h(x,y) = \begin{cases} 0, & \text{if } x = y, \\ xy, & \text{otherwise.} \end{cases}$$

Similarly, the Axioms (Diss1) and (Diss2) are fulfilled. We continue with Axiom (Div3). The same partition from the previous example $X = X_1 \cup X_2 \cup X_3 \cup X_4 \cup X_5 \cup X_6$ will be considered. Then

$$D(A,B) = \sum_{i=1}^{6} \sum_{x \in X_i} h(A(x), B(x)).$$

Applying the standard minimum t-norm we have:

$$D(A \cap C, B \cap C) = \sum_{x \in X_1 \cup X_3} h(A(x), B(x)) + \sum_{x \in X_2} h(A(x), C(x))$$

$$+ \sum_{x \in X_4} h(C(x), B(x)) + \sum_{x \in X_5 \cup X_6} h(C(x), C(x)).$$

Since h is positive and increasing in both components, we have

- $h(A(x), B(x)) = h(A(x), B(x))$ for all $x \in X_1 \cup X_3$,
- $h(A(x), C(x)) < h(A(x), B(x))$ for all $x \in X_2$,
- $h(C(x), B(x)) < h(A(x), B(x))$ for all $x \in X_4$,
- $h(C(x), C(x)) = 0 \leq h(A(x), B(x))$ for all $x \in X_5 \cup X_6$.

Finally,

$$D(A \cap C, B \cap C) \leq \sum_{x \in X} h(A(x), B(x)) = D(A,B).$$

We have shown that $D(A \cap C, B \cap C) \leq D(A,B)$. However, the inequality for the union is not fulfilled in general. Thus, if we consider the universal set formed by only one element, $X = \{x\}$, the dual maximum t-conorm S_M and the fuzzy sets A, B, C defined again as $A(x) = 0.2, B(x) = 0.8, C(x) = 0.5$, we have $D(A \cup C, B \cup C) = h((A \cup C)(x), (B \cup C)(x)) = h(S_M(A(x), C(x)), S_M(B(x), C(x))) = h(C(x), B(x)) = 0.5 \cdot 0.8 = 0.4$ and $D(A,B) = h(A(x), B(x)) = 0.2 \cdot 0.8 = 0.16$, and therefore $D(A \cup C, B \cup C) > D(A,B)$.

Thus, the inequality for the intersection $D(A \cap C, B \cap C) \leq D(A,B)$ does not imply the inequality for the union $D(A \cup C, B \cup C) \leq D(A,B)$, and vice versa.

From the previous examples we can conclude that both conditions in Axiom 3 of Definition 19 are independent, and therefore both are necessary to define a divergence measure.

7 Relation between divergences and dissimilarities

As mentioned in the previous section, divergence measures appeared as an alternative to dissimilarities. We could think that both concept are related in general, but it is not true. Thus, we will show one example of divergence measure which is not a dissimilarity and conversely, we will consider a dissimilarity which is not a divergence measure.

Example 26. Let us consider the divergence measure D defined on the one-point referential set $X = \{x\}$ as follows:

$$D(A, B) = \begin{cases} 0, & \text{if } A = B \text{ or } A(x) \in \{0,1\} \text{ or } B(x) \in \{0,1\}, \\ 1, & \text{otherwise.} \end{cases}$$

Now, we verify that D is a divergence measure. The first and the second conditions are trivial. Let us have a look at the third one:

(i) If $A = B$, then $A \cap C = B \cap C$ and $A \cup C = B \cup C$ for any $C \in \mathcal{F}(X)$. Thus, $D(A \cap C, B \cap C) = D(A \cup C, B \cup C) = D(A, B) = 0$.

(ii) If $A \neq B$ and $A(x), B(x) \in (0, 1)$, then $D(A, B) = 1$ by definition and since $D(A \cap C, B \cap C), D(A \cup C, B \cup C) \in \{0, 1\}$ by definition, they are less than or equal to $D(A, B)$.

(iii) If $A = \varnothing$ and $B \neq \varnothing$, then $D(A, B) = 0$ by definition.
For the intersection, we have that $(A \cap C)(x) = T(A(x), C(x)) = T(0, C(x)) = 0$, and therefore $A \cap C = \varnothing$. Thus, $D(A \cap C, B \cap C) = 0 \leq D(A, B)$.
For the union, we have that $(A \cup C)(x) = S(A(x), C(x)) = S(0, C(x)) = C(x)$, and therefore $A \cup C = C$. Thus,

- if $C(x) \in \{0, 1\}$, then $D(A \cup C, B \cup C) = D(C, B \cup C) = 0$;
- if $C(x) \in (0, 1)$, then $B \cup C = X$ if we consider the union defined by means of the drastic t-conorm. Thus, $D(A \cup C, B \cup C) = D(C, X) = 0$.

In all the cases, $D(A \cup C, B \cup C) \leq D(A, B)$.

(iv) The case $B = \varnothing$ and $A \neq \varnothing$ can be proved analogously to the case (iii).

(v) If $A = X$ and $B \neq X$, then $D(A, B) = 0$ by definition. Moreover, $A \cap C = C$ and $A \cup C = X$, since 1 is the neutral element of any t-norm and the annihilating element of any t-conorm. Thus, the divergence between the intersections assumes the value zero for any B if we consider the drastic t-norm, since

- if $C(x) \in \{0, 1\}$, then $D(A \cap C, B \cap C) = D(C, B \cap C) = 0$ by definition;
- if $C(x) \in (0, 1)$, then $B \cap C = \varnothing$ if we consider T_D to define the intersection. Thus, $D(A \cap C, B \cap C) = D(C, \varnothing) = 0$.

In both cases, $D(A \cap C, B \cap C) \leq D(A, B)$. Moreover, $D(A \cup C, B \cup C) = D(X, B \cup C) = 0 \leq D(A, B)$ from the definition of the divergence D.

(vi) The case $B = X$ and $A \neq X$ can be proved analogously to the case (v).

Thus, the map D is a divergence. However, if we consider the fuzzy sets A, B, C whose membership functions are $A(x) = 0, B(x) = 0.3, C(x) = 0.5$, we have that $A \subseteq B \subseteq C$ but $D(B, C) = 1 > 0 = D(A, C)$. Thus, D does not fulfill Axiom (Diss.3) in Definition 18.

Thus not all divergences are dissimilarities. The converse does not hold either as we prove by the following example.

Example 27. Let $D : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ be defined by:

$$D(A, B) = \begin{cases} 0, & \text{if } A = B, \\ 1, & \text{if } A \neq B \text{ and either } A = X \text{ or } B = X, \\ 0.5, & \text{otherwise.} \end{cases}$$

Let us check that D is a dissimilarity. The first and second conditions are trivial. Consider now the fuzzy sets $A, B, C \in \mathcal{F}(X)$ for which $A \subseteq B \subseteq C$. We are going to prove that $D(A,C) \geq \max\{D(A,B), D(B,C)\}$. Now $D(A,C)$ only can attain three values:

- if $D(A,C) = 1$, then $D(A,B), D(B,C) \in \{0, 0.5, 1\} \leq 1 = D(A,B)$.
- if $D(A,C) = 0.5$, then $A \neq C, A \neq X, C \neq X$. Since $A \subseteq B \subseteq C$ we take $B \neq X$ and hence $D(A,B), D(B,C) \in \{0, 0.5\} \leq 0.5 = D(A,C)$.
- if $D(A,C) = 0$, then $A = C$ by definition. Since $A \subseteq B \subseteq C$ we have $A = B = C$ and hence $D(A,B) = D(B,C) = 0 \leq D(A,B)$.

Thus, D is a dissimilarity, but it is not a divergence. Consider the fuzzy sets A, B, C defined on two-point universe $X = \{x_1, x_2\}$ as follows: $A(x_1) = 1, A(x_2) = 0, B(x_1) = 0, B(x_2) = 1, C = B$. Thus, in this case we have that $D(A \cup C, B \cup C) = D(X, B) = 1 > 0.5 = D(A,B)$ for any t-conorm. Therefore, (Div.3) from Definition 19 is not satisfied.

We conclude that divergences and dissimilarities are not related in general, a divergence need not be a dissimilarity and vice versa. However, there are some maps which belong to both families, as we can see in the next example.

Example 28. For the map from Example 20, we proved that D is a divergence measure for any t-norm and t-conorm. Moreover, if $A \subseteq B \subseteq C$, we have:

- $D(A,C) = 1 \Rightarrow D(A,B), D(B,C) \in \{0,1\} \leq D(A,B)$,
- $D(A,C) = 0 \Rightarrow A = C \Rightarrow A = B = C \Rightarrow D(A,B) = D(B,C) = 0$.

Thus, it is also a dissimilarity.

In the following proposition a relation between a dissimilarity and a divergence measure associated to (X, T_M, S_M) is discussed.

Proposition 29. *Let D be a divergence measure associated to the triple (X, T_M, S_M). Then the map D is a dissimilarity measure.*

Proof. The first and the second conditions fulfilled by any divergence and dissimilarity, respectively, are the same. To prove the third one we consider three fuzzy subsets A, B, C, for which $A \subseteq B \subseteq C$. Since D is a divergence measure, we have the following:

- $D(A,B) = D(A \cap B, C \cap B) \leq D(A,C)$,
- $D(B,C) = D(A \cup B, C \cup B) \leq D(A,C)$,

and therefore $D(A,C) \geq \max\{D(A,B), D(B,C)\}$. Thus, D is a dissimilarity measure. □

Although in general both concept are not related, they are some functions which are both divergences and dissimilarities. In fact, if in the particular case we reduce our focus on the standard fuzzy set operations using only minimum t-norm and its dual t-conorm, then we can consider the divergences as a subset of dissimilarities (see the previous Proposition 29 and [14]), that is,

$$\text{Divergence} \overset{(X,T_M,S_M)}{\Longrightarrow} \text{Dissimilarity}.$$

Their relationship is even stronger for a particular kind of families, which have the local property, as we will see in the next section.

8 Distance-based divergence measures

Our aim is to extend the previous approach to measure a difference of two fuzzy subsets discussed in [14] and [15] with the new one based on cardinality. Motivation comes from real life situations in which we want to compare two real objects on the basis of required criteria. Divergence measure given by suitable distance is well-founded, however, in the same situation it does not give us a complete information.

Let $A, B \in \mathcal{F}(X)$. We assume the following membership degrees of elements $x, y \in X : A(x) = 0.47, B(x) = 0.57, A(y) = 0.52, B(y) = 0.62$. We pose the following question: What is a standard distance between the values $A(x), B(x)$ and $A(y), B(y)$? The answer should be 0.10 in both cases. For certain reasons we can say that the elements having the membership degrees less than 0.5 are not significant. The definition of scalar cardinality allows us to modify previous results by means of a function f as follows: we put $f(A(x)) = 0$ if $A(x) < 0.5$. Now, we see that the elements $x, y \in X$ become more different in this sense.

In the succeeding paragraph we define a distance measure for fuzzy sets.

Definition 30. A map $d : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ is a distance measure if and only if the function d satisfies the following conditions:

(1) for all $A, B \in \mathcal{F}(X); d(A, B) \geq 0$ and $d(A, B) = 0$ if and only if $A = B$;

(2) for all $A, B \in \mathcal{F}(X); d(A, B) = d(B, A)$;

(3) for all $A, B, C \in \mathcal{F}(X); d(A, C) \leq d(A, B) + d(B, C)$.

The divergences and distances are not related in general. Examples of a distance measure which is not a divergence measure and vice versa, will be shown in Example 31 and Example 32. More general examples can be found in [10].

Example 31. Consider the map $d : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ defined by

$$d(A, B) = \begin{cases} 0, & \text{if } A = B, \\ 1, & \text{if } A \neq B, \text{ but } A = X \text{ or } B = X, \\ 0.5, & \text{otherwise.} \end{cases}$$

Let us prove that d is a distance measure. Positivity, the identity of indiscernibles and symmetry hold trivially. It remains to show that it also has the triangular inequality property. Let $A, B, C \in \mathcal{F}(X)$.

- If $d(A, C) = 0$, the inequality holds trivially.
- If $d(A, C) = 0.5$, then $A \neq C$, and therefore either $B \neq A$ or $B \neq C$, and consequently $d(A, C) = 0.5 \leq d(A, B) + d(B, C)$.
- Finally, if $d(A, C) = 1$, we can assume, without loss of generality, that $A = X$. If $B = A$, then $d(A, B) = 0$ and $d(B, C) = 1$, and consequently, $d(A, C) = 1 = d(A, B) + d(B, C)$. Otherwise, if $B \neq A$, then $d(A, B) = 1$, and therefore, $d(A, C) = 1 \leq d(A, B) + d(B, C)$.

We conclude that $d(A, C) \leq d(A, B) + d(B, C)$ in all cases. Thus, the map d is a distance measure. It is also a dissimilarity measure by Example 27.

However, the map d is not a divergence measure as we have seen in Example 27.

We will continue with an example of a divergence measure which is not a distance measure.

Example 32. Consider the map $D : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ defined in the following way:

$$D(A,B) = \max_{x \in X} \left(A(x) - B(x)\right)^2,$$

where $A, B \in \mathcal{F}(X)$ and X is a finite universe. In the case that X is infinite, a supremum instead of a maximum is necessary to consider.

In the first step, we will prove that D associated to (X, T_M, S_M) is a divergence. It is easy to see that $D(A,B) = 0$ if and only if $A = B$ and $D(A,B) = D(B,A)$. Next, without loss of generality, we suppose $A(x) \geq B(x)$ for some $x \in X$. Then, for the fuzzy set C three cases are discussed:

- if $C(x) > A(x) \geq B(x)$, then $|\min(A(x), C(x)) - \min(B(x), C(x))| = |A(x) - B(x)|$,
- if $A(x) \geq C(x) \geq B(x)$, then $|\min(A(x), C(x)) - \min(B(x), C(x))| = |C(x) - B(x)| \leq |A(x) - B(x)|$,
- if $A(x) \geq B(x) > C(x)$, then $|\min(A(x), C(x)) - \min(B(x), C(x))| = |C(x) - C(x)| = 0 \leq |A(x) - B(x)|$.

For each $x \in X$ we have $|T_M(A(x), C(x)) - T_M(B(x), C(x))| \leq |A(x) - B(x)|$, and therefore:

$$\left(T_M(A(x), C(x)) - T_M(B(x), C(x))\right)^2 \leq \left(A(x) - B(x)\right)^2.$$

Applying the previous relation to all elements $x \in X$ we obtain:

$$\max_{x \in X} \left(T_M(A(x), C(x)) - T_M(B(x), C(x))\right)^2 \leq \max_{x \in X} \left(A(x) - B(x)\right)^2,$$

i.e.

$$D(A \cap_{T_M} C, B \cap_{T_M} C) \leq D(A,B).$$

Analogously it can be proved that $D(A \cup_{S_M} C, B \cup_{S_M} C) \leq D(A,B)$. We conclude that the map D is a divergence measure.

In the second step, we show that the map D is not a distance. We consider three fuzzy sets A, B, C defined on the one-point universe $X = \{x\}$, such that $A(x) = 0$, $B(x) = 0.3$, $C(x) = 0.7$. Then we have:

$$D(A,C) = 0.49 > 0.25 = 0.09 + 0.16 = D(A,B) + D(B,C),$$

and therefore the triangle inequality is not fulfilled. So, the map D is not a distance measure.

Let $d : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ be a distance measure and $A, B \in \mathcal{F}(X)$. Then we will define pointwise a distance measure for each $x \in X$ as follows:

$$d(A,B) = d_0(A(x), B(x)),$$

where d_0 defined on $[0,1] \times [0,1]$ is a restricted distance of the original one d.

Proposition 33. *A mapping* $|\cdot| : \mathcal{F}(X) \to [0, \infty[$ *is a scalar cardinality if for each* $A \in \mathcal{F}(X)$*:*

$$|A| = \sum_{x \in Supp(A)} f\left(A(x)\right),$$

where $f : [0,1] \to [0,1]$ *is a function for which the following conditions are fulfilled:*

(a) boundary conditions: $f(0) = 0, f(1) = 1$*;*

(b) monotonicity: for all $a, b \in [0,1] : a \leq b \Rightarrow f(a) \leq f(b)$*.*

In the case of crisp finite sets, the definition of cardinality leads to the classical formulation of the “number of elements” (having the membership values equal to 1). We say that the divergence measure D in the next proposition is generated by the distance d_0.

Proposition 34. *Let* $A, B \in \mathcal{F}(X)$ *and* $|\cdot|$ *denotes a scalar cardinality. Let the map* $D : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ *be defined in the following way:*

$$D(A,B) = \left|\Phi\left(A,B\right)\right|,$$

where $\Phi : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathcal{F}(X)$ *is a function defined axiomatically as:*

$$A(x) \times B(x) \overset{\Phi}{\longmapsto} C(x),$$

in which $C(x) = d_0(A(x), B(x))$ *and* $d_0 : [0,1] \times [0,1] \to [0,1]$ *is a suitable distance (see Remark 35). Then the map* D *is a divergence measure between fuzzy sets* A *and* B*.*

Proof. Let us check that D is a divergence measure. If $A = B$, then $\Phi(A(x), B(x)) = 0$ and $f(0) = 0$, and therefore the divergence between A and B is zero. A reverse implication is not requested. It is evident that D is also commutative, since:

$$D(A,B) = |\Phi(A,B)| = \sum_{x \in X} f\left(\Phi(A(x), B(x))\right) =$$

$$= \sum_{x \in X} f\left(\Phi(B(x), A(x))\right) = |\Phi(B,A)| = D(B,A).$$

If the fuzzy subsets A and B become more similar in the sense of union or intersection, then the value of divergence will decrease and the inequalities $D(A \cup C, B \cup C) \leq D(A,B), D(A \cap C, B \cap C) \leq D(A,B)$ are fulfilled since we have used only a special class of distances (see Remark 35) and monotonicity of f from Proposition 33. □

Remark 35. It is necessary to clarify the notion of a suitable distance. As we have seen in Example 31, the set of distances and divergences are not comparable in general since neither concept of these can imply the other one. The distances which are also the divergences, are very important for our work and only this class of distances will be considered in further text. Both of the divergence measures introduced in Example 36 and Example 37 are also the distance measures.

Example 36. A simple example of a distance is the function $d(x, y) = |x - y|$. Then the following inequalities are fulfilled for $T \in \{T_M, T_P, T_L\}$ and $S \in \{S_M, S_P, S_L\}$: $d(T(x, z), T(y, z)) \leq d(x, y)$ and $d(S(x, z), S(y, z)) \leq d(x, y)$. It follows immediately from Example 21. We conclude that d is a divergence measure.

Example 37. Let us define the function d in the following way:

$$d(x, y) = \begin{cases} c, & \text{if } x \neq y, \text{ where } 0 < c \leq 1, \\ 0, & \text{if } x = y. \end{cases}$$

It is easy to see that d is a distance. We are going to show that d is also a divergence. The first and the second condition hold trivially. To prove the third one, we can consider two cases:

- If $d(x, y) = c$, then immediately $d(T(x, z), T(y, z)) \in \{0, c\} \leq c = d(x, y)$.
- If $d(x, y) = 0$, then $x = y$, and therefore $T(x, z) = T(y, z)$ and $d(T(x, z), T(y, z)) = 0 = d(x, y)$.

We have shown that $d(T(x, z), T(y, z)) \leq d(x, y)$. Similarly, it can be proved that $d(S(x, z), S(y, z)) \leq d(x, y)$. We conclude that d is a divergence, too, although this example is quite an extreme case.

We have introduced the class of distance-based divergence measures. It is obvious that all distances cannot be considered and a utility of the concept of distance-based cardinalities must be restricted to a special class of distances.

9 Local divergence measures

When we compare two fuzzy sets, it seems natural to suppose that if we only change the value of these sets at one element, the divergence should only depend on what has been changed.

Definition 38. Let D be a **divergence measure** for a triple (X, T, S). Then D has the local property or, briefly, is **local**, if for all $A, B \in \mathcal{F}(X)$ and for all $x \in X$, there exists a map $h_x : [0, 1] \times [0, 1] \to \mathbb{R}$ such that

$$D(A, B) - D(A \cup \{x\}, B \cup \{x\}) = h_x(A(x), B(x)).$$

A particular case of locality was introduced in [12, 14], but there h_x was fixed for any $x \in X$ and all the elements in the universe were of the same importance. However, based on some application of comparison of multivalued sets (see, e.g. [20, 21]), this is not always the case and different maps should be considered. Therefore we introduce a more general definition for locality.

In the case X is a finite universe, a representation theorem for local divergences was obtained in [14]. However, this result holds only for the minimum t-norm and its dual t-norm, that is, in (X, T_M, S_M). We will present a general result which holds for any t-norm and any t-conorm.

Theorem 39 (Representation Theorem 1)**.** *Let (X, T, S) be a triple with X a finite universe and T and S any t-norm and t-conorm, respectively. Let D be a divergence associated to (X, T, S). D is local if and only if*

$$D(A, B) = \sum_{x \in X} h_x(A(x), B(x)),$$

where $\{h_x\}_{x\in X}$ *is a family of maps from* $[0,1]\times[0,1]$ *into* $\mathbb{R}$ *such that for any* $x \in X$ *and* $a, b, c \in [0,1]$ *the following hold:*

(i) $h_x(a,a) = 0$*, for all* $a \in [0,1]$*,*

(ii) $h_x(a,b) = h_x(b,a)$*, for all* $a, b \in [0,1]$*,*

(iii) $h_x(a,b) \geq \max(h_x(S(a,c),S(b,c)), h_x(T(a,c),T(b,c)))$ *for all* $a,b,c \in [0,1]$*.*

Proof. If D is local, then, by definition, $D(A,B) - D(A\cup\{x\}, B\cup\{x\}) = h_x(A(x),B(x))$ for any $x \in X$. We apply this equation recursively for other elements from the referential X. Therefore:

$$D(A,B) = D(A\cup\{x\}, B\cup\{x\}) + h_x(A(x),B(x)) = D(A\cup\{x\}\cup\{y\}, B\cup\{x\}\cup\{y\}) + h_x(A(x),B(x)) + h_y((A\cup\{x\})(y),(B\cup\{x\})(y)).$$

But,

$$(A\cup\{x\})(y) = S(A(y),\{x\}(y)) = S(A(y),0) = A(y),$$

and analoguously,

$$(B\cup\{x\})(y) = B(y).$$

Thus,

$$\begin{aligned} D(A,B) &= D(A\cup\{x\}\cup\{y\}, B\cup\{x\}\cup\{y\}) + h_x(A(x),B(x)) + h_y(A(y),B(y)) \\ &= D(A\cup X, B\cup X) + \sum_{x\in X} h_x(A(x),B(x)) = D(X,X) + \sum_{x\in X} h_x(A(x),B(x)) \\ &= \sum_{x\in X} h_x(A(x),B(x)), \end{aligned}$$

since D is a divergence measure and $D(X,X) = 0$.

We have to check if for any $x \in X$, the function $h_x : [0,1]\times[0,1] \to \mathbb{R}$ fulfills the properties (i)-(iii). For any fixed $x \in X$, we can define the fuzzy sets $A(t) = a, B(t) = b, C(t) = c$ if $t = x$, and $A(t) = B(t) = C(t) = 0$ otherwise.

(i) In the above case we have $0 = D(A,A) = \sum_{x\in X} h_x(A(x),A(x)) = h_x(a,a)$.

(ii) By definition we have that $D(A,B) = D(B,A)$, but $D(A,B) = h_x(a,b)$ and $D(B,A) = h_x(b,a)$.

(iii) As D is a divergence measure we have $D(A,B) \geq \max(D(A\cup C, B\cup C), D(A\cap C, B\cap C))$, but $D(A,B) = h_x(a,b)$ and

$$D(A\cup C, B\cup C) = \sum_{x\in X} h_x(S(A(x),C(x)), S(B(x),C(x))) =$$

$$= h_x(S(a,c),S(b,c)) + \sum_{t\in X-\{x\}} h_t(S(0,0),S(0,0)) = h_x(S(a,c),S(b,c))$$

by property (i) previously proved. Analogously we could prove that $D(A\cap C, B\cap C) = h_x(T(a,c),T(b,c))$, and therefore the proof of this implication is concluded.

Let us show the converse implication. Let $D : \mathcal{F}(X) \times \mathcal{F}(X) \to \mathbb{R}$ be a map, which can be expressed as a sum

$$D(A,B) = \sum_{x\in X} h_x(A(x), B(x)),$$

where $\{h_x\}_{x\in X}$ is a family of maps from $[0,1] \times [0,1]$ into $\mathbb{R}$ such that, for any $x \in X$, h_x fulfills the conditions (i)-(iii) from the previous definition. Now, we will show that D is a local divergence measure. It is immediate that D is well-defined. Apart from that, D is a divergence, since for any A, B, C in $\mathcal{F}(X)$ we have:

(Diss.1)
$$D(A,A) = \sum_{x\in X} h_x(A(x), A(x)) = 0,$$

(Diss.2)
$$D(A,B) = \sum_{x\in X} h_x(A(x), B(x)) = \sum_{x\in X} h_x(B(x), A(x)) = D(B,A),$$

(Div.3)
$$D(A\cap C, B\cap C) = \sum_{x\in X} h_x(T(A(x), C(x)), T(B(x), C(x)))$$
$$\leq \sum_{x\in X} h_x(A(x), B(x)) = D(A,B),$$

and analogously, $D(A \cup C, B \cup C) \leq D(A,B)$.

Moreover,

$$D(A,B) - D(A\cup\{x\}, B\cup\{x\}) = \sum_{t\in X} h_t(A(t), B(t)) - \sum_{t\in X} h_t((A\cup\{x\})(t), (B\cup\{x\})(t))$$
$$= \sum_{t\in X-\{x\}} (h_t(A(t), B(t)) - h_t(S(A(t),0), S(B(t),0)))$$
$$+(h_x(A(x), B(x)) - h_x(S(A(t),1), S(B(t),1)))$$
$$= \sum_{t\in X-\{x\}} (h_t(A(t), B(t)) - h_t(A(t), B(t))) + (h_x(A(x), B(x)) - h_x(1,1)) = h_x(A(x), B(x))$$

for any $A, B \in \mathcal{F}(X)$ and any $x \in X$. Thus, D is a local divergence measure. □

An example of a local divergence measure is now presented.

Example 40. The divergence proposed in Example 21 is local. It is clear that for any $x \in X$, $h_x(a,b) = \alpha_x \cdot |a-b|$ for all $a, b \in [0,1]$ where $\alpha_x \geq 0$ for any $x \in X$ and $\sum_{x\in X} \alpha_x = 1$. Thus, it suffices to prove that it fulfills the conditions in Theorem 39. The first two are trivial. For the third one the proof follows immediately from Example 21.

In the previous example, the map h_x was based on a distance between real numbers. However, this is not true in general (see [15]). Only some specific distances can be used for generating divergence measures, in fact, h_x is itself a divergence measure on the referential $\{x\}$.

Although the divergence measure given in Example 21 is local, this is not true in general. It is clear that there exist divergence measures which are not local. As an example we can consider the measure proposed in Example 20. Suppose that D is local, then there is a family of maps $\{h_x\}_{x\in X}$ with the appropriate properties such that $D(A,B) = \sum_{x\in X} h_x(A(x), B(x))$ for any $A, B \in \mathcal{F}(X)$.

In particular, if we consider three fuzzy sets A, B, C such that $A(x) = a_1$ and $A(y) = a_2$ for two fixed points $x, y \in X$,

$$B(t) = \begin{cases} b, & \text{if } t = x \\ A(x), & \text{otherwise} \end{cases} \quad \text{and} \quad C(t) = \begin{cases} c, & \text{if } t = y \\ A(x), & \text{otherwise} \end{cases}$$

with $a_1 \neq b$ and $a_2 \neq c$, then we have, by definition of D, that $1 = D(A,B) = h(a_1, b)$ and $1 = D(A,C) = h(a_2, c)$.

On the other hand, the fuzzy sets B, C are also different and $D(B,C) = 1$. But if D is local, we also have that $D(B,C) = h(b, a_1) + h(a_2, c) = 1 + 1 = 2 \neq 1$, which contradicts the definition of D. Thus, we conclude that D is a divergence measure, but it has not the local property.

Thus, the family of local divergences is a proper subset of the family of divergences. This subfamily has some specific properties as we will see in the following results. Moreover, these properties will be very important in the next section, when we generalize the concept of locality.

Apart from the definition for local divergences, we can also consider a different one. The main advantage of this new definition is that it allows us to generalize this notion later.

Proposition 41. *Let (X, T, S) be a triple with X a finite universe and T and S any t-norm and t-conorm, respectively. Let D be a divergence associated to (X, T, S). A divergence D has the local property if and only if*

$$D(A,B) = D(A \cup Z, B \cup Z) + D(A \cup Z^c, B \cup Z^c)$$

for any $A, B \in \mathcal{F}(X)$ and any $Z \in \mathcal{P}(X)$.

Proof. If D is local, then there exists an appropriate family of functions $\{h_x\}_{x\in X}$ such that, for any $A, B \in \mathcal{F}(X)$ and any $Z \in \mathcal{P}(X)$,

$$D(A,B) = \sum_{x\in Z} h_x(A(x), B(x)) + \sum_{x\in Z^c} h_x(A(x), B(x)),$$

$$D(A \cup Z, B \cup Z) = \sum_{x\in Z} h_x(1,1) + \sum_{x\in Z^c} h_x(A(x), B(x)) = \sum_{x\in Z^c} h_x(A(x), B(x))$$

and

$$D(A \cup Z^c, B \cup Z^c) = \sum_{x\in Z} h_x(A(x), B(x)) + \sum_{x\in Z^c} h_x(1,1) = \sum_{x\in Z} h_x(A(x), B(x))$$

by the properties of t-norms and t-conorms. So, $D(A,B) = D(A \cup Z, B \cup Z) + D(A \cup Z^c, B \cup Z^c)$.

Conversely, if D fulfills $D(A,B) = D(A \cup Z, B \cup Z) + D(A \cup Z^c, B \cup Z^c)$ for any $A, B \in \mathcal{F}(X)$ and any $Z \in \mathcal{P}(X)$, for each $x \in X$ we can consider $Z = \{x\}$ and therefore $D(A,B) - D(A \cup \{x\}, B \cup \{x\}) = D(A \cup \{x\}^c, B \cup \{x\}^c)$. Thus, we only have to prove that $h_x(u,v) = D(A_{x,u}, A_{x,v})$ fulfills the conditions in Theorem 39 for any $x \in X$, where $A_{x,u}(x) = u$ and $A_{x,u}(y) = 1$ for any $y \neq x$.

The first two conditions are easy:

$$h_x(u,u) = D(A_{x,u}, A_{x,u}) = 0$$

and

$$h_x(u,v) = D(A_{x,u}, A_{x,v}) = D(A_{x,v}, A_{x,u}) = h_x(v,u).$$

For the third one, let us consider any $u, v, z \in [0,1]$. Then

$$\begin{aligned} h_x(u,v) &= D(A_{x,u}, A_{x,v}) \\ &\geq \max(D(A_{x,u} \cup A_{x,z}, A_{x,v} \cup A_{x,z})), D(A_{x,u} \cap A_{x,z}, A_{x,v} \cap A_{x,z}) \\ &= \max(D(A_{x,S(u,z)}, A_{x,S(v,z)}), D(A_{x,T(u,z)}, A_{x,T(v,z)})) \\ &= \max(h_x(S(u,z), S(v,z)), h_x(T(u,z), T(v,z))). \end{aligned}$$

Therefore, D is the local divergence from Theorem 39. □

Once we have equivalent definitions for local divergences, we can prove two properties of them in the following propositions.

Proposition 42. *Let (X,T,S) be a triple with X a finite universe and let $S(A)$ be the support of the fuzzy subset A. Let D be a local divergence associated to X. Let A, B and C be three fuzzy subsets with $S(B) = S(C)$ and $S(A) \cap S(B) = \varnothing$. Then $D(A \cup B, A \cup C) = D(S(A) \cup B, S(A) \cup C)$.*

Proof. By the previous proposition we have that $D(A \cup B, A \cup C) = D(A \cup B \cup S(A), A \cup C \cup S(A)) + D(A \cup B \cup S(A)^c, A \cup C \cup S(A)^c)$. But it is easy to prove that $A \cup B \cup S(A) = B \cup S(A)$, $A \cup C \cup S(A) = C \cup S(A)$ and $A \cup B \cup S(A)^c = A \cup C \cup S(A)^c$. Thus, $D(A \cup B, A \cup C) = D(S(A) \cup B, S(A) \cup C)$. □

Corollary 43. *Let (X,T,S) be a triple with X a finite universe. Let D be a local divergence associated to X. Let $A, A', B, B' \in \mathcal{F}(X)$. If there exists $Z \in \mathcal{P}(X)$ such that $D(A \cup Z, B \cup Z) \leq D(A' \cup Z, B' \cup Z)$ and $D(A \cup Z^c, B \cup Z^c) \leq D(A' \cup Z^c, B' \cup Z^c)$, then $D(A,B) \leq D(A',B')$.*

It is natural to ask if the distance-based divergence has a local property or not. The result is obtained in the following proposition.

Proposition 44. *Let D be a divergence measure which is generated by the distance d_0. Then D has the local property.*

Proof. The divergence D can be rewritten by Proposition 33 and Proposition 34 as follows:

$$D(A,B) = |\Phi(A,B)| = \sum_{x \in X} f\left(\Phi(A(x), B(x))\right) = \sum_{x \in X} f\left(d_0(A(x), B(x))\right).$$

We can see that the divergence measure D can be expressed pointwise as a sum of suitable distances d_0. Each such distance d_0 satisfies all conditions for the function h

from Definition 38 and properties of local divergence given in Theorem 39. To show it, the distance d_0 takes the value 0 if both coordinates are the same. The map d_0 is also symmetric. The third axiom of divergence is fulfilled since we have considered only a restricted class of divergence-generating distances d_0. Applying the function f from Proposition 33, all properties of the function h (see Representation Theorem 39) are kept without any change. More formally, for each $x \in X$ the following conditions hold:

(1) $f(\Phi(A(x), A(x))) = f(0) = 0$,

(2) $f(\Phi(A(x), B(x))) = f(\Phi(B(x), A(x)))$,

(3) $f(\Phi((A\cup C)(x), (B\cup C)(x))) \leq f(\Phi(A(x), B(x)))$ and $f(\Phi((A\cap C)(x), (B\cap C)(x))) \leq f(\Phi(A(x), B(x)))$.

Then we can define $h(a, b) = f(\Phi(a, b))$ for all $a, b \in [0, 1]$. It allows us to write:

$$D(A, B) = |\Phi(A, B)| = \sum_{x\in X} h(A(x), B(x)).$$

Thus, the divergence D has the local property. □

Example 45. Let $X = \{x_1, x_2, x_3\}$ and $A, B \in \mathcal{F}(X)$. Consider the following membership degrees: $A(x_1) = 0.2$, $A(x_2) = 0.8$, $A(x_3) = 0.7$, $B(x_1) = 0.3$, $B(x_2) = 0.3$, $B(x_3) = 0.3$ and the functions $f(x) = x;\ d_0(x, y) = |x - y|$. Then:

$$D(A, B) = |\Phi(A, B)| = \sum_{x\in X} f\left(d_0(A(x), B(x))\right) = \sum_{x\in X} |A(x)-B(x)| = 0.1+0.5+0.4 = 1.0.$$

10 Local dissimilarities

The previous properties have been proved in the general case of any t-norm and t-conorm. If in the particular case we consider the standard union or intersection, S_M and T_M, an interesting property can be proved. It will allow us to relate divergences and dissimilarities. First we characterize the dissimilarities with the local property.

Theorem 46 (Representation Theorem 2)**.** *Let (X, T, S) be a triple with X a finite universe and T and S any t-norm and t-conorm, respectively. Let D be a* ***dissimilarity measure*** *associated to (X, T, S). Then D is* ***local****, that is, for all $A, B \in \mathcal{F}(X)$ and for all $x \in X$, there exists a map $h_x : [0, 1] \times [0, 1] \to \mathbb{R}$ such that*

$$D(A, B) - D(A \cup \{x\}, B \cup \{x\}) = h_x(A(x), B(x))$$

if and only if

$$D(A, B) \ = \ \sum_{x\in X} h_x(A(x), B(x)),$$

where, for any $x \in X$, h_x is a function with the following properties:

(i) $h_x(a, a) = 0$, *for all* $a \in [0, 1]$;

(ii) $h_x(a, b) = h_x(b, a)$, *for all* $a, b \in [0, 1]$;

(iii') $h_x(a, c) \geq \max\{h_x(a, b), h_x(b, c)\}$, *for all* $a, b, c \in [0, 1]$ *with* $a < b < c$.

Proof. Let D be a dissimilarity with the local property. From locality we can prove that

$$D(A,B) \;=\; \sum_{x\in X} h_x(A(x),B(x))$$

and as the first and the second axioms of dissimilarity are the same as the axioms for divergences, we proved in Theorem 39 that h_x fulfills that $h_x(a,a)=0$ and $h_x(a,b)=h_x(b,a)$, for all $a,b\in[0,1]$.

Moreover, for any $a,b,c\in[0,1]$ such that $a\le b\le c$, if we apply Axiom (Diss.3) in Definition 18 to A,B,C with $A(x)=a$, $B(x)=b$, $C(x)=c$ for a fixed $x\in X$ and $A(y)=B(y)=C(y)=0$ if $y\neq x$, then we have that $h_x(a,c)=D(A,C)\ge \max(D(A,B),D(B,C))=\max(h_x(a,b),h_x(b,c))$.

The converse implication is analogous to the proof for local divergences (Theorem 39). □

In the following result we will study the relationship between the locality of divergence and locality of dissimilarity in the particular case of the minimum t-norm and its dual t-conorm.

Proposition 47. *Let (X,T_M,S_M) be a triple with X a finite universe and T_M and S_M the minimum t-norm and the maximum t-conorm, respectively. A map $D:\mathcal{F}(X)\times\mathcal{F}(X)\to\mathbb{R}$ is a local divergence if only if D is a local dissimilarity.*

Proof. It is sufficient to prove that (iii) in Theorem 39 is equivalent to (iii') in Theorem 46.

If (iii) is fulfilled and $a\le b\le c$, then $h_x(b,c)=h_x(S_M(a,b),S_M(c,b))\le h_x(a,c)$ and $h_x(a,b)=h_x(T_M(a,b),T_M(c,b))\le h_x(a,c)$.

Conversely, consider without loss of generality that $a\le b$. We discuss three cases:

- If $c\le a\le b$, then $h_x(S_M(a,c),S_M(b,c))=h_x(a,b)$ and $h_x(T_M(a,c),T_M(b,c))=h_x(c,c)=0\le h_x(a,b)$.
- If $a\le c\le b$, then $h_x(S_M(a,c),S_M(b,c))=h_x(c,b)\le h_x(a,b)$ and $h_x(T_M(a,c),T_M(b,c))=h_x(a,c)\le h_x(a,b)$, by applying condition (iii') in both cases.
- If $a\le b\le c$, then $h_x(S_M(a,c),S_M(b,c))=h_x(c,c)=0\le h_x(a,b)$ and $h_x(T_M(a,c),T_M(b,c))=h_x(a,b)$.

□

Thus, although divergences and dissimilarities are different concepts, in the case of the minimum t-norm and its dual t-conorm, any divergence is a dissimilarity. Now, we have proved that the family of the local divergences is the same as the family of the local dissimilarities, that is,

$$\{\text{Local divergences }\} \overset{(X,T_M,S_M)}{=} \{\text{ Local dissimilarities}\}.$$

As we could see in Proposition 22, any divergence measure has 0 as a lower bound. Moreover, if it is local and we consider T_M and S_M, it also fulfills Axiom (Diss.3) of

Definition 18. Thus, $D(A', B') \leq D(\varnothing, X)$ for any $A' \subset B'$. Moreover, for arbitrary $A, B \in \mathcal{F}(X)$, if D is local we also have that $D(A, B) = D(A', B')$ with

$$A'(x) = \begin{cases} A(x) & \text{if } A(x) \leq B(x), \\ B(x) & \text{otherwise,} \end{cases} \qquad B'(x) = \begin{cases} B(x) & \text{if } A(x) \leq B(x), \\ A(x) & \text{otherwise.} \end{cases}$$

Thus, $D(A, B) \leq D(\varnothing, X)$, for any $A, B \in \mathcal{F}(X)$, that is, $D(\varnothing, X)$ is an upper bound for D.

From now on, we can consider the normalized version of a divergence measure, that is, $\frac{1}{D(\varnothing,X)} \cdot D$. Let us notice that the normalized version assumes values in the interval $[0, 1]$.

11 Generalized local divergences

As we could see, the class of local divergence measures has important properties. So, it is interesting to consider a wider class of divergences with these properties. This is the purpose of this section.

As we can see in Proposition 41, a divergence D is local if and only if $D(A, B) = D(A \cup Z, B \cup Z) + D(A \cup Z^c, B \cup Z^c)$, for any $A, B \in \mathcal{F}(X)$ and any $Z \in \mathcal{P}(X)$. In this case we are decomposing by means of the sum, but any other aggregation function (see, for instance, [4]) could be considered. Thus, we could consider the particular family of divergence measures such that

$$D(A, B) = Ag(D(A \cup Z, B \cup Z), D(A \cup Z^c, B \cup Z^c)),$$

for any $A, B \in \mathcal{F}(X)$ and any $Z \in \mathcal{P}(X)$ and a fixed aggregation function Ag.

Let us study the behaviour of this map Ag, in detail:

- Since it is fulfilled for any crisp set, it is fulfilled in particular for any crisp set Z and for its complement Z^c. Thus, $D(A, B) = Ag(D(A \cup Z, B \cup Z), D(A \cup Z^c, B \cup Z^c))$ and $D(A, B) = Ag(D(A \cup Z^c, B \cup Z^c), D(A \cup Z, B \cup Z))$ and therefore Ag has to be commutative.
- If we consider as the crisp set the referential X, then $D(A, B) = Ag(D(A \cup X, B \cup X), D(A \cup \varnothing, B \cup \varnothing)) = Ag(D(X, X), D(A, B)) = Ag(0, D(A, B))$. Thus, zero has to be the neutral element for Ag.
- We see that for the local divergences, if $D(A \cup Z, B \cup Z) \leq D(A' \cup Z, B' \cup Z)$ and $D(A \cup Z^c, B \cup Z^c) \leq D(A' \cup Z^c, B' \cup Z^c)$ for a crisp set Z, then $D(A, B) \leq D(A', B')$. If we require this natural property in our general context, then Ag has to be increasing in both components.
- Moreover, we can apply Ag recursively and obtain that $D(A, B) = Ag(D(A \cup Z, B \cup Z), D(A \cup Z^c, B \cup Z^c)) = Ag(Ag(D(A \cup Z \cup Z, B \cup Z \cup Z), D(A \cup Z \cup Z^c, B \cup Z \cup Z^c)), D(A \cup Z^c, B \cup Z^c)) = Ag(Ag(D(A \cup Z \cup Z^c, B \cup Z \cup Z^c), D(A \cup Z, B \cup Z)), D(A \cup Z^c, B \cup Z^c))$, but also $D(A, B) = Ag(D(A \cup X, B \cup X), D(A \cup \varnothing, B \cup \varnothing)) = Ag(D(A \cup Z \cup Z^c, B \cup Z \cup Z^c), Ag(D(A \cup Z, B \cup Z), D(A \cup Z^c, B \cup Z^c)))$.
  Thus, Ag has to be associative.

Finally, Ag has to be an associative, commutative, monotone function with neutral element 0, that is, Ag has to be a t-conorm.

The previous comments allow us to generalize the notion of locality as follows.

Definition 48. Let $D : \mathcal{F}(X) \times \mathcal{F}(X) \to [0,1]$ be a **divergence measure** and let S be a t-conorm. It is said that D has the **S-local property** or, briefly, that it is S-local, if for all $A, B \in \mathcal{F}(X)$ and for all $Z \in \mathcal{P}(X)$ and $x \in X$, the following property is fulfilled:

$$D(A,B) = S[D(A \cup Z, B \cup Z), D(A \cup Z^c, B \cup Z^c)].$$

Again it is possible to characterize this family of divergences.

Theorem 49 (Representation Theorem 3)**.** *Let (X, T, S) be a triple with X a finite universe and T and S any t-norm and t-conorm, respectively. Let S' be another t-conorm. A map D is an S'-local divergence measure if and only if*

$$D(A,B) \;=\; \underset{x \in X}{S'} \left[h_x(A(x), B(x))\right],$$

where $\{h_x\}_{x \in X}$ is a family of maps such that for any $x \in X$, $h_x : [0,1] \times [0,1] \to [0,1]$ and

(i) $h_x(a,a) = 0$, *for all* $a \in [0,1]$;
(ii) $h_x(a,b) = h_x(b,a)$, *for all* $a, b \in [0,1]$;
(iii) $h_x(a,b) \geq \max(h_x(S(a,c), S(b,c)), h_x(T(a,c), T(b,c)))$, *for all* $a, b, c \in [0,1]$.

Proof. In this theorem we apply recursively the t-conorm. Thus, it means that we have $S'(x,y,z) = S'(x, S'(y,z))$, etc.

If D is S'-local, then

$$D(A,B) = S'(D(A \cup \{x_1\}, B \cup \{x_1\}), D(A \cup \{x_1\}^c, B \cup \{x_1\}^c))$$

$$= S'(D(A \cup \{x_1\}, B \cup \{x_1\}), S'(D(A \cup \{x_1\}^c \cup \{x_2\}, B \cup \{x_1\}^c \cup \{x_2\}), D(A \cup \{x_1\}^c \cup \{x_2\}^c, B \cup \{x_1\}^c \cup \{x_2\}^c)))$$

$$= S'(D(A \cup \{x_1\}, B \cup \{x_1\}), S'(D(A \cup \{x_2\}, B \cup \{x_2\}), D(X,X))) = \cdots$$

$$\cdots = S'\left(\underset{x \in X}{S'}\left[D(A \cup \{x\}, B \cup \{x\}), D(X,X)\right]\right) = S'\left(\underset{x \in X}{S'}\left[D(A \cup \{x\}, B \cup \{x\}), 0\right]\right)$$

$$= \underset{x \in X}{S'}\left[D(A \cup \{x\}, B \cup \{x\})\right].$$

Thus, it is enough to prove that for any $x \in X$, the map h_x defined by $h_x(u,v) = D(A_{x,u}, A_{x,v})$ is well-defined and it fulfills the conditions (i)–(iii) for any $x \in X$, where $A_{x,u}(x) = 1$ and $A_{x,u}(y) = u$ for any $y \neq x$. The proof is analogical to the one given in Proposition 41.

It is easy to prove that the map D defined in the statement is a divergence measure. For the S'-locality,

$$D(A,B) = S'\left(\underset{x \in Z}{S'}\left[h_x(A(x), B(x))\right], \underset{x \in Z^c}{S'}\left[h_x(A(x), B(x))\right]\right)$$

and

$$D(A\cup Z, B\cup Z) = \underset{x\in Z}{S'}\left[h_x((A\cup Z)(x),(B\cup Z)(x))\right] + \underset{x\in Z^c}{S'}\left[h_x((A\cup Z)(x),(B\cup Z)(x))\right] =$$

$$= \underset{x\in Z}{S'}[0] + \underset{x\in Z^c}{S'}\left[h_x(A(x),B(x))\right] = \underset{x\in Z^c}{S'}\left[h_x(A(x),B(x))\right].$$

Analogously,

$$D(A\cup Z^c, B\cup Z^c) = \underset{x\in Z}{S'}\left[h_x(A(x),B(x))\right].$$

Thus,

$$D(A,B) = S'\left(D(A\cup Z, B\cup Z), D(A\cup Z^c, B\cup Z^c)\right)$$

for any $Z\in\mathcal{P}(X)$, that is, D is a S'-local divergence measure. □

From the previous theorem it is immediate that any divergence measure assuming values in the interval $[0,1]$ is local if and only if it is S_L-local.

As this holds for local divergences, an equivalent definition will be given by means of the intersection instead of the union in the following proposition.

Proposition 50. *A divergence measure D has the S-local property if and only if for all $A,B\in\mathcal{F}(X)$ and for all $Z\in\mathcal{P}(X)$ the following equality is fulfilled:*

$$D(A,B) = S\left(D(A\cap Z, B\cap Z), D(A\cap Z^c, B\cap Z^c)\right)$$

Proof. If D is S-local, by the properties of the t-conorms and the t-norms we have that

$$S\left(D(A\cap Z, B\cap Z), D(A\cap Z^c, B\cap Z^c)\right)$$

$$= \underset{x\in Z}{S}\left[S\left(h((A\cap Z)(x),(B\cap Z)(x)), h((A\cap Z^c)(x),(B\cap Z^c)(x))\right)\right]$$

$$+ \underset{x\in Z^c}{S}\left[S\left(h((A\cap Z)(x),(B\cap Z)(x)), h((A\cap Z^c)(x),(B\cap Z^c)(x))\right)\right]$$

$$= \underset{x\in Z}{S}\left[S\left(h(A(x),B(x)), h(0,0)\right)\right] + \underset{x\in Z^c}{S}\left[S\left(h(0,0), h(A(x),B(x))\right)\right] = D(A,B).$$

Conversely, if D is a divergence measure by fulfilling the condition for the intersection, it is easy to prove that D can be decomposed as in Theorem 49 for the map h_x defined by $h_x(u,v) = D(A_{x,u}, A_{x,v})$, where $A_{x,u}(x) = u$ and $A_{x,u}(y) = 1$ for any $y\neq x$. Thus, D is S-local. □

Another equivalent definition for S-locality is shown below.

Proposition 51. *Let D be a divergence measure and let S be a t-conorm. D is S-local if and only if for any $A,B\in\mathcal{F}(X)$ and for any crisp partitions $\{Z_i\}_{i=1}^m$ of the reference space X, the following equation holds*

$$D(A,B) = \overset{m}{\underset{i=1}{S}}\left[D(A\cap Z_i, B\cap Z_i)\right].$$

Proof. If D is S-local, by the previous proposition we have that

$$D(A,B) = S\left(D(A\cap Z_1, B\cap Z_1), D(A\cap Z_1^c, B\cap Z_1^c)\right).$$

But if we apply again the result to the second component, we have that

$$D(A\cap Z_1^c, B\cap Z_1^c) = S\left(D(A\cap Z_1^c\cap Z_2, B\cap Z_1^c\cap Z_2), D(A\cap Z_1^c\cap Z_2^c, B\cap Z_1^c\cap Z_2^c)\right).$$

Since $A\cap Z_1^c\cap Z_2 = A\cap Z_2$ and $B\cap Z_1^c\cap Z_2 = B\cap Z_2$, and considering the associativity of the t-conorm, we obtain

$$D(A,B) = S\left(\mathop{S}_{i=1}^{2}\left[D(A\cap Z_i, B\cap Z_i), D(A\cap Z_1^c\cap Z_2^c, B\cap Z_1^c\cap Z_2^c)\right]\right).$$

If we apply the previous procedure recursively, we obtain that

$$D(A,B) = S\left(\mathop{S}_{i=1}^{m}\left[D(A\cap Z_i, B\cap Z_i), D(A\cap Z_1^c\cap\ldots\cap Z_m^c, B\cap Z_1^c\cap\ldots\cap Z_m^c)\right]\right) =$$

$$= S\left(\mathop{S}_{i=1}^{m}\left[D(A\cap Z_i, B\cap Z_i), D(\varnothing,\varnothing)\right]\right) = S\left(\mathop{S}_{i=1}^{m}\left[D(A\cap Z_i, B\cap Z_i), 0\right]\right) =$$

$$= \mathop{S}_{i=1}^{m}\left[D(A\cap Z_i, B\cap Z_i)\right].$$

The converse implication is trivial, since we only have to consider for any $Z\in\mathcal{P}(X)$ the partition $\{Z, Z^c\}$ of X and Proposition 50. □

In the following propositions we establish some important properties of the S-local divergences, which express natural characteristics of our measure.

The first one expresses that the divergence between any crisp set and its complement is constant.

Proposition 52. *Let D be an S-local divergence. Then $D(Z,Z^c) = D(X,\varnothing)$ for all $Z\in\mathcal{P}(X)$.*

Proof. Let Z be any crisp subset of X. Then:

$$D(Z,Z^c) = \mathop{S}_{x\in X}\left[h_x(Z(x),Z^c(x))\right] = S\left(\mathop{S}_{x\in Z}\left[h_x(Z(x),Z^c(x))\right], \mathop{S}_{x\in Z^c}\left[h_x(Z(x),Z^c(x))\right]\right)$$

$$= S\left(\mathop{S}_{x\in Z}\left[h_x(1,0)\right], \mathop{S}_{x\in Z^c}\left[h_x(0,1)\right]\right) = \mathop{S}_{x\in X}\left[h_x(1,0)\right] = D(X,\varnothing),$$

where we have applied the symmetry of h_x for any $x\in X$ and the associativity of the t-conorm. □

Although trivial, the following proposition allows us to change the scale factor of a divergence according to our particular requirements.

Proposition 53. *Let D be a S-local divergence generated by the family of functions $\{h_x\}_{x\in X}$, and let $\phi:[0,1]\to[0,1]$ be an increasing function with $\phi(0)=0$. The maps D_ϕ and D^ϕ defined by*

$$D_\phi(A,B) = S_{x\in X}\left[\phi(h(A(x),B(x))\right],$$
$$D^\phi(A,B) = \phi\left(S_{x\in X}\left[h(A(x),B(x))\right]\right),$$

are also divergence measures and D_ϕ is S-local.

Proof. For D^{ϕ}, it is trivial that $(Diss.2)$ is fulfilled. Since $\phi(0) = 0$, the same holds for $(Diss.1)$. The monotonicity of ϕ implies that $(Div.3)$ is also fulfilled and therefore D^{ϕ} is a divergence measure.

For proving that D_{ϕ} is an S-local divergence measure, we only have to prove that the family $\{\phi \circ h_x\}_{x \in X}$ fulfills conditions in Theorem 49. It is again immediate from the properties of ϕ. □

As we have seen in the previous section, the particular case of the classical t-norm and t-conorm has interesting properties. We can obtain specific results also in this case, which are shown below.

Proposition 54. *Let (X, T_M, S_M) be a triple with X a finite universe, let S be any t-conorm and let D be an S-local divergence measure. For any $A, B, C \in \mathcal{F}(X)$ such that for any $x \in X$ either $A(x) \leq B(x) \leq C(x)$ or $C(x) \leq B(x) \leq A(x)$, we have that*

$$D(A, C) \geq \max\left(D(A, B), D(B, C)\right).$$

Proof. Let us consider $Z = \{x \in X \mid A(x) \leq B(x) \leq C(x)\}$. Then, by the statement, we have that $Z^c \subseteq \{x \in X \mid C(x) \leq B(x) \leq A(x)\}$. Thus,

$D(A, B) = S\left(S_{x \in Z}\left[h_x(A(x), B(x))\right], S_{x \in Z^c}\left[h_x(A(x), B(x))\right]\right)$.

But for any $x \in Z$ we have that $A(x) = T_M(A(x), B(x))$ and $B(x) = T_M(C(x), B(x))$ and for any $x \in Z^c$ we have that $A(x) = S_M(A(x), B(x))$ and $B(x) = S_M(C(x), B(x))$.

Thus,

$$D(A, B) =$$

$$S\left(\underset{x \in Z}{S}\left[h_x(T_M(A(x), B(x)), T_M(C(x), B(x)))\right], \underset{x \in Z^c}{S}\left[h_x(S_M(A(x), B(x)), S_M(C(x), B(x)))\right]\right).$$

But, by (iii) in Theorem 49 we have that

$$h_x(T_M(A(x), B(x)), T_M(C(x), B(x))) \leq h_x(A(x), C(x))$$

and

$$h_x(S_M(A(x), B(x)), S_M(C(x), B(x))) \leq h_x(A(x), C(x)).$$

Thus,

$$D(A, B) \leq D(A, C).$$

Analogously, we can prove that $D(B, C) \leq D(A, C)$. □

Now, it is immediate to prove the following result.

Corollary 55. *Let (X, T_M, S_M) be a triple with X a finite universe, let S be any t-conorm and let D be an S-local divergence measure. For any $A_1, A_2, B_1, B_2 \in \mathcal{F}(X)$ such that $A_1 \subseteq B_1 \subseteq B_2 \subseteq A_2$, we have $D(B_1, B_2) \leq D(A_1, A_2)$.*

12 Applications

In the previous sections we have presented a theoretical study of comparison measures for fuzzy sets. But, as we commented in the introduction, the comparison of the fuzzy sets is important in many fields. We will present two examples of application of generalized local divergence measures for pattern recognition and decision making. In both cases we will remark the different situations depending on the selected t-conorm used to decompose the divergence.

12.1 Applications to pattern recognition

We suppose that m patterns and also the sample are represented by the fuzzy sets. Our goal is to choose some pattern, which differs from the sample at the least possible grade (in sense of divergence measure). In the examples we use generalized local divergences and further we compare some particular results for each basic t-conorm. In order to obtain more representative results we can consider a weight vector, which assigns a particular importance to each element of the universal set.

Let X be a finite universe, let assume that the patterns $A_1, A_2, ..., A_m$ are represented by fuzzy sets and let B be a sample represented also by a fuzzy set.

As we can measure the difference between B and A_i for $i \in \{1, ..., m\}$, we obtain the finite set of divergences: $D(A_1, B), \ldots, D(A_m, B)$.

Finally, the sample B will be associated to the pattern A_j whenever

$$D(A_j, B) = \min_{i=1,\ldots,m} D(A_i, B).$$

That means, the sample B is classified into the pattern from which it differs least.

The following example is based on the one proposed in [20].

Example 56. Let us consider five kinds of mineral fields, each of them featured by the content of six minerals and containing one kind of typical hybrid mineral. Those five kinds of typical hybrid mineral are represented by fuzzy sets A_1, A_2, A_3, A_4 and A_5 in $X = \{x_1, \ldots, x_6\}$, respectively. Let us assume that there is another kind of hybrid mineral B, and we want to classify it into one of the aforementioned mineral fields. The minerals are described by means of the fuzzy sets defined in Table 1.

X	x_1	x_2	x_3	x_4	x_5	x_6
A_1	0.74	0.03	0.19	0.49	0.02	0.74
A_2	0.12	0.03	0.05	0.14	0.02	0.39
A_3	0.45	0.66	1.00	1.00	1.00	1.00
A_4	0.28	0.52	0.47	0.30	0.19	0.74
A_5	0.33	1.00	0.18	0.16	0.05	0.68
B	0.63	0.52	0.21	0.22	0.07	0.66

Table 1. The kinds of hybrid minerals represented by fuzzy sets.

Let us also assume that experts have established the following weight vector on X as follows: $\alpha = \{0.2, 0.3, 0.125, 0.125, 0.125, 0.125\}$. We will use our method to classify B. We consider the divergence measure proposed in Examples 21 and 40:

$$D(A, B) = \underset{x \in X}{S} \left[\alpha_x \cdot |A(x) - B(x)|\right]$$

with $\alpha_x \geq 0$ for any $x \in X$ and $\sum_{x \in X} \alpha_x = 1$. Then, $S_{L\, x \in X}\, \alpha_x = 1 \geq S_{P\, x \in X}\, \alpha_x \geq S_{M\, x \in X}\, \alpha_x$, therefore it is clear that D is a S-local divergence measure for the maximum,

	S_M	S_P	S_L	S_D
$D(A_1, B)$	0.1470	0.2090	0.2215	1
$D(A_2, B)$	0.1470	0.2864	0.3190	1
$D(A_3, B)$	0.1163	0.3644	0.4330	1
$D(A_4, B)$	0.0700	0.1314	0.1375	1
$D(A_5, B)$	0.1440	0.2084	0.2203	1

Table 2. The S-local divergences obtained for $S \in \{S_M, S_P, S_L, S_D\}$.

the product and the Łukasiewicz t-conorms. The same happens for the drastic t-conorm. Thus, the values assumed by the four different divergences are obtained in Table 2.

We can see that in all the cases we should classify B into the hybrid mineral A_4 (in the case of the drastic t-conorm it is just an option). However, the behaviour of any divergence is different. Thus, for instance, for the maximum t-conorm, the divergence is not able to distinguish between A_1 and A_2, since in this case only the maximum distance is considered. In the other two cases, S_P and S_L, all the points in the referential X are essential to obtain the value of the divergence and therefore the difference between A_1 and A_2 is remarked. In the case of S_D, the information given by the divergence is insignificant.

Apart from the t-conorm used to define the S-local divergence, the weights can also play an interesting role. Thus, at the previous example, we will consider the weight vector $\alpha = \{k, 0.5 - k, 0.125, 0.125, 0.125, 0.125\}$ for $k \in \{0, 0.1, \ldots, 0.5\}$.

In this case, if we consider the previous divergence measure D for S_M, S_P and S_L, respectively, we obtain the following results:

(1) Results for S_M-local divergences are shown in Table 3:

	$k=0$	$k=0.1$	$k=0.2$	$k=0.3$	$k=0.4$	$k=0.5$
$D(A_1, B)$	0.2450	0.1960	0.1470	0.0980	0.0490	0.0550
$D(A_2, B)$	0.2450	0.1960	0.1470	0.1530	0.2040	0.2550
$D(A_3, B)$	0.1163	0.1163	0.1163	0.1163	0.1163	0.1163
$D(A_4, B)$	0.0325	0.0350	0.0700	0.1050	0.1400	0.1750
$D(A_5, B)$	0.2400	0.1920	0.1440	0.0960	0.1200	0.1500

Table 3. Variations of the S_M-local divergence based on the weights.

The following relations are fulfilled:
$D(A_4, B) < D(A_3, B) < D(A_5, B) < D(A_1, B) = D(A_2, B)$, for $k \in \{0, 0.1, 0.2\}$,
$D(A_5, B) < D(A_1, B) < D(A_4, B) < D(A_3, B) < D(A_2, B)$, for $k = 0.3$,
$D(A_1, B) < D(A_3, B) < D(A_5, B) < D(A_4, B) < D(A_2, B)$, for $k \in \{0.4, 0.5\}$.

In all cases, A_2 is the least appropriate hybrid mineral. Furthermore, a behaviour of other patterns depends on the parameter k and the weight vector α. Whereas the sample B should be classified into the pattern A_4 for $k \leq 0.2$, an importance of the pattern A_4 for greater k is decreasing. On the contrary, the opposite relation for the pattern A_1 is fulfilled.

We conclude that the sample B should be classified into the pattern:

- A_1 for $k = 0.4, 0.5$,
- A_4 for $k = 0, 0.1, 0.2$,
- A_5 for $k = 0.3$.

(2) Results for S_P-local divergences are shown in Table 4:

	$k=0$	$k=0.1$	$k=0.2$	$k=0.3$	$k=0.4$	$k=0.5$
$D(A_1, B)$	0.2841	0.2460	0.2090	0.1729	0.1379	0.1039
$D(A_2, B)$	0.2966	0.2892	0.2864	0.2883	0.2948	0.3060
$D(A_3, B)$	0.3599	0.3620	0.3644	0.3671	0.3702	0.3737
$D(A_4, B)$	0.0660	0.0987	0.1314	0.1640	0.1967	0.2294
$D(A_5, B)$	0.2523	0.2289	0.2084	0.1907	0.1758	0.1637

Table 4. Variations of the S_P-local divergence based on the weights.

The following relations are fulfilled:
$D(A_4, B) < D(A_5, B) < D(A_1, B) < D(A_2, B) = D(A_3, B)$, for $k \in \{0, 0.1, 0.2\}$,
$D(A_4, B) < D(A_1, B) < D(A_5, B) < D(A_2, B) < D(A_3, B)$, for $k = 0.3$,
$D(A_1, B) < D(A_5, B) < D(A_4, B) < D(A_2, B) < D(A_3, B)$, for $k \in \{0.4, 0.5\}$.

In all cases, A_3 is the least appropriate hybrid mineral to be considered. The behaviour of the patterns A_1, A_4 is similar as in the previous case, in which S_M-local divergence has been considered. As we can see, one basic difference in results of pattern recognition compared to S_M-local divergences for $k = 0.3$ is remarkable, where the pattern B should be classified into A_4 instead of A_5.

We conclude that the sample B should be classified into the pattern:

- A_1 for $k = 0.4, 0.5$,
- A_4 for $k = 0, 0.1, 0.2, 0.3$.

(3) Results for S_L-local divergences are shown in Table 5:

The following relations are fulfilled:
$D(A_4, B) < D(A_5, B) < D(A_1, B) < D(A_2, B) = D(A_3, B)$, for $k \in \{0, 0.1, 0.2\}$,
$D(A_4, B) < D(A_1, B) < D(A_5, B) < D(A_2, B) < D(A_3, B)$, for $k = 0.3$,
$D(A_1, B) < D(A_5, B) < D(A_4, B) < D(A_2, B) < D(A_3, B)$, for $k \in \{0.4, 0.5\}$.

The behaviour in this case is the same as we have remarked for S_P-local divergences. Only insignificant numerical differences are observed.

We conclude that the sample B should be classified into the pattern:

	$k=0$	$k=0.1$	$k=0.2$	$k=0.3$	$k=0.4$	$k=0.5$
$D(A_1,B)$	0.2975	0.2595	0.2215	0.1835	0.1455	0.1075
$D(A_2,B)$	0.3150	0.3170	0.3190	0.3210	0.3230	0.3250
$D(A_3,B)$	0.4250	0.4290	0.4330	0.4370	0.4410	0.4450
$D(A_4,B)$	0.0675	0.1025	0.1375	0.1725	0.2075	0.2425
$D(A_5,B)$	0.2563	0.2383	0.2203	0.2023	0.1843	0.1663

Table 5. Variations of the S_L-local divergence based on the weights.

- A_1 for $k=0.4, 0.5$,
- A_4 for $k=0, 0.1, 0.2, 0.3$.

The obtained results are totally expected, since for $k=0.5$ the most important point at the referential is x_1. If we look at this element, A_1 and B are the most similar. In the case $k=0$, the most important point changes to x_2 and this is the reason of the selection of A_4.

As we have seen in the previous examples, if we consider different maps h_x for any $x \in X$ when defining the divergence measure, we can increase the information obtained from the data. The same happens if we consider the concept of generalized locality, since we can aggregate the divergences among the points by different t-conorm apart from S_L (just the sum) and we can consider the most appropriate t-conorm to any problem. Thus, the information obtained by means of the class of divergence measures considered here is much richer than the one considered in [14].

12.2 Applications to decision making

Now, we will apply previous theoretical results in the multiple attribute decision making. We consider m alternatives, all of them represented by the fuzzy sets, the set of n attributes and the associated weight vector α. The new alternatives are considered - as "the most appropriate" and as "the least appropriate". In this sense the preferred alternative should have as large as possible degree of similarity with the most optimal one and should have as large as possible degree of difference with the least optimal one. We show how an alternative responding to these requirements can be chosen. A calculation for basic triangular conorms is presented, obtained results are compared and final results are scheduled in tables for particular S-local divergences. The choice of the preferred alternative can be influenced also by using of the weight vector associated to a particular problem.

First we present the following notation:
let $A = \{A_1, \ldots, A_m\}$ denote a set of m alternatives,
let $X = \{x_1, \ldots, x_n\}$ be a set of n attributes,
and $\alpha = (\alpha_1, \ldots, \alpha_n)$ be its associated weight vector, where $\alpha_i \geq 0$ and $\sum_i \alpha_i = 1$.

Each alternative A_i will be expressed by a fuzzy set with the elements x_j, where $A_i(x_j)$ represents the degree in which alternative A_i agrees with attribute x_j. We create the new alternatives A^+ and A^- defined by

$$A^+ = \bigcup_{i=1}^{m} A_i, \qquad \text{and} \qquad A^- = \bigcap_{i=1}^{m} A_i.$$

The alternatives A^+ and A^- can be interpreted as the "optimal" and the "least optimal", respectively. In this sense, the preferred alternative A would be more similar to A^+ and more different from A^-, simultaneously.

Finally, we consider the quotient k_i defined as:

$$k_i = \frac{D(A_i, A^+)}{D(A_i, A^+) + D(A_i, A^-)}.$$

It means that if some alternative A_j has a quotient k_j for which $k_j < k_i$ for all $i \neq j$, then the alternative A_j is better than A_i in the sense previously described. Thus, the optimal is the alternative A_i whose k_i is the minimum.

The previous procedure will now be explained by the means of an example based on the one proposed in [21].

Example 57. The government has to decide among five different energy strategies: $A_1 - A_5$. Each of them is assessing four attributes: economic (x_{EC}), technological (x_T), environmental (x_{EN}) and socio-political (x_P). The following weight vector of these attributes $(\alpha_{EC}, \alpha_T, \alpha_{EN}, \alpha_P) = (0.4, 0.2, 0.3, 0.1)$ will be considered.

Let us assume that alternatives A_i are defined by the fuzzy sets given in Table 6.

X	x_{EC}	x_T	x_{EN}	x_P
A_1	0.2	0.7	0.6	0.5
A_2	0.4	0.5	0.8	0.6
A_3	0.5	0.6	0.9	0.7
A_4	0.3	0.8	0.7	0.5
A_5	0.8	0.7	0.1	0.3

Table 6. Definition of 5 energy strategies.

We will consider the triple (X, T, S), where S is used to define the union and T to define the intersection of five alternatives represented by fuzzy sets. The corresponding fuzzy sets A^+ and A^- defined in Table 7 will be computed. For all basic t-norms and t-conorms, the obtained results will be compared. The results are illustrated in the following four cases:

(1) for the triple (X, T_M, S_M) and S_M-local divergence we have:

If we consider the S_M-local divergence measure proposed in the previous example with the maximum t-conorm, that is,

$$D(A, B) = \underset{x \in X}{S_M} \left[\alpha_x \cdot |A(x) - B(x)|\right],$$

X	x_{EC}	x_T	x_{EN}	x_P
A^+	0.8	0.8	0.9	0.7
A^-	0.2	0.5	0.1	0.3

Table 7. Definition of the most optimal and the least optimal alternatives for (X, T_M, S_M).

	$D(A^+, A_i)$	$D(A^-, A_i)$	k_i
$i = 1$	0.24	0.15	0.62
$i = 2$	0.16	0.21	0.43
$i = 3$	0.12	0.24	0.33
$i = 4$	0.20	0.18	0.53
$i = 5$	0.24	0.24	0.50

Table 8. Comparison of optimality of five alternatives for S_M-local divergence.

then the following divergences and quotients will be obtained:

Since $k_3 < k_2 < k_5 < k_4 < k_1$, we conclude that in accordance with the considered criteria the most optimal alternative is A_3.

(2) for the triple (X, T_P, S_P) and S_P-local divergence we have:

X	x_{EC}	x_T	x_{EN}	x_P
A^+	0.9664	0.9964	0.9978	0.9790
A^-	0.0096	0.1176	0.0302	0.0315

Table 9. Definition of the most optimal and the least optimal alternatives for (X, T_P, S_P).

Now, if we consider the S_P-local divergence measure with the probabilistic sum, that is,

$$D(A, B) = \underset{x \in X}{S_P} \left[\alpha_x \cdot |A(x) - B(x)|\right],$$

then the following results will be obtained:

	$D(A^+, A_i)$	$D(A^-, A_i)$	k_i
$i = 1$	0.4531	0.3550	0.5607
$i = 2$	0.3695	0.4348	0.4594
$i = 3$	0.2933	0.4992	0.3701
$i = 4$	0.3891	0.4188	0.4816
$i = 5$	0.4020	0.4244	0.4864

Table 10. Comparison of optimality of five alternatives for S_P-local divergence.

Since $k_3 < k_2 < k_4 < k_5 < k_1$, we conclude that the most optimal alternative is again A_3. However, the behaviour differs from previous example. Since $k_5 < k_4$ for S_M-local divergence, then $k_4 < k_5$ for S_P-local divergence.

(3) for the triple (X, T_L, S_L) and S_L-local divergence we have:

X	x_{EC}	x_T	x_{EN}	x_P
A^+	1	1	1	1
A^-	0	0	0	0

Table 11. Definition of the most optimal and the least optimal alternatives for (X, T_L, S_L).

If we consider the S_L-local divergence measure proposed in the previous example with the Łukasiewicz t-conorm, that is,

$$D(A, B) = \underset{x \in X}{S_L} \left[\alpha_x \cdot |A(x) - B(x)|\right],$$

then we obtain the following results:

Since $k_3 < k_2 < k_4 < k_5 < k_1$, the same result as we have seen in the case of S_P-divergences can be concluded.

	$D(A^+, A_i)$	$D(A^-, A_i)$	k_i
$i = 1$	0.55	0.45	0.55
$i = 2$	0.44	0.56	0.44
$i = 3$	0.34	0.66	0.34
$i = 4$	0.46	0.54	0.46
$i = 5$	0.48	0.52	0.48

Table 12. Comparison of optimality of five alternatives for S_L-local divergence.

(4) for the triple (X, T_D, S_D) and S_D-local divergence we have:

X	x_{EC}	x_T	x_{EN}	x_P
A^+	1	1	1	1
A^-	0	0	0	0

Table 13. Definition of the most optimal and the least optimal alternatives for (X, T_D, S_D).

If we consider the S_D-local divergence measure proposed in the previous example with the drastic sum, that is,

$$D(A, B) = \underset{x \in X}{S_D} \left[\alpha_x \cdot |A(x) - B(x)| \right],$$

then the following divergences and quotients will be obtained:

	$D(A^+, A_i)$	$D(A^-, A_i)$	k_i
$i = 1$	1	1	0.5
$i = 2$	1	1	0.5
$i = 3$	1	1	0.5
$i = 4$	1	1	0.5
$i = 5$	1	1	0.5

Table 14. Comparison of optimality of five alternatives for S_D-local divergence.

Since $k_i = 0.5$ for all $i \in \{1, \ldots, 5\}$, one can not make decision based on the obtained information.

We could consider not only different t-conorms, but also different weights for any particular problem, and again different results could be obtained.

13 Conclusion and open problems

We have introduced and justified a new class of divergence measures between two fuzzy subsets, the S-local divergences, which are constructed by using a triangular conorm S

instead of the sum. In these studies we have used the previous results about divergence measures with the local property, in fact we have shown that any local divergence is an S-local divergence. Since this new class of t-conorms dependent divergences satisfies all the properties already proposed for the local ones, we think that this generalization will allow us to manage a larger class of divergence measures and use them to define new classes of fuzziness measures.

One of the advantages of the proposed attitude is that by considering different maps h_x for the points (or ranges of points) in X, we can stress the specific aspects of a particular model, and thus obtain more accurate information when specifying the measure of difference.

Therefore, we have some immediate open problems for the future research. The most tantalizing one is to generalize the S-local divergence measure to the case when the universe is infinite. Another important point is the choice of h_x in some typical cases.

Finally, we have discussed some applications of S-local divergences for pattern recognition and decision making. The other generalizations of these concepts can be considered, e.g. using family of weight vectors presented on richer sample. We have restricted this research to only four basic t-conorms, but any other can be taken into account. Next, some applications for further areas, for instance, fuzzy control systems, can be considered.

References

[1] J. Aczél, "Lectures on Functional Equations and their Applications", Academic Press, New York, 1966.

[2] D. Blizard, The Development of Multiset Theory, *Modern Logic* **1** (1991), 319–352.

[3] B. Bouchon-Meunier, M. Rifqi, S. Bothorel, Towards general measures of comparison of objects, *Fuzzy Sets and Systems* **84** (1996), 143–153.

[4] T. Calvo, G. Mayor, R. Mesiar, "Aggregation operators: New trends and Applications", Physica-Verlag, 2002.

[5] I. Couso, L. Garrido, L. Sánchez, Similarity and dissimilarity measures between fuzzy sets: A formal relational study, *Information Sciences* **229** (2013), 122–141.

[6] E. P. Klement, R. Mesiar, E. Pap, "Triangular Norms", Kluwer Academic Publishers, London, 2000.

[7] V. Kobza, V. Janiš, S. Montes, Generalizated local divergence measures, *Journal of Intelligent and Fuzzy Systems* **33** (2017), 337–350.

[8] V. Kobza, V. Janiš, S. Montes, Divergence measures on hesitant fuzzy sets, *Journal of Intelligent and Fuzzy Systems* **33** (2017), 1589–1601.

[9] C. M. Ling, Representation of associative functions, *Publ. Math. Debrecen* **12** (1965), 189–212.

[10] I. Montes, "Comparison of alternatives under uncertainty and imprecision", Ph.D. Dissertation, University of Oviedo, Oviedo, Spain, 2014.

[11] I. Montes, V. Janiš, N. R. Pal, S. Montes, Local Divergences for Atanassov Intuitionistic Fuzzy Sets, *IEEE Transactions on Fuzzy Systems* **24** (2015), 360–373.

[12] S. Montes, "Partitions and divergence measures in fuzzy models", Ph.D. Dissertation, University of Oviedo, Oviedo, Spain, 1998.

[13] S. Montes, I. Couso, C. Bertoluzza, Some classes of fuzziness measures from local divergences, *Belgian Journal of Operations Research, Statistics and Computer Science* **38** (1998), 37–49.

[14] S. Montes, I. Couso, C. Bertoluzza, Fuzziness measures from divergence measures, *Proceedings of the conference EUROFUSE-SIC'99, Budapest* (1999), 365–370.

[15] S. Montes, I. Couso, P. Gil, C. Bertoluzza, Divergence measure between fuzzy sets, *International Journal of Approximate Reasoning* **30** (2002), 91–105.

[16] S. Montes, P. Gil, Some classes of divergence measures between fuzzy subsets and between fuzzy partitions, *Mathware and Soft Computing* **5** (1998), 253–265.

[17] B. Schweizer, A. Sklar, Statistical metric spaces, *Pacific J. Math* **10** (1960), 313–334.

[18] B. Schweizer, A. Sklar, Associative functions and statistical triangle inequalities, *Publ. Math. Debrecen*, **10** (1963), 69–81.

[19] B. Schweizer, A. Sklar, "Probabilistic metric spaces", North-Holland, New York, 1983.

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

[21] Z. Xu, M. Xia, Distance and similarity measures for hesitant fuzzy sets, *Information Sciences* **181** (2011), 2128–2138.

[22] L. Zadeh, Fuzzy sets, *Information and Control* **8** (1965), 338–353.

[23] L. Zadeh, "Fuzzy sets and systems", System Theory, Polytechnic Press, Brooklyn, 1965.