Towards a mathematical framework for similarity and dissimilarity

Jorge Orozco and Lluís Belanche

Grupo de Soft-Computing
http://www.lsi.upc.es/dept/investigacion/sectia/soco
Departament de Sistemes i Llenguatges Informàtics
Universitat Politècnica de Catalunya
Mòdul C6 - Campus Nord. Jordi Girona, E-08034 Barcelona, España
{jorozco,belanche}@lsi.upc.es

Abstract. This work contributes to the design and understanding of similarity and dissimilarity in AI, in order to increase their general utility. A formal definition for each concept is proposed, joined with a set of fundamental properties. A main basis of results are compiled by application of transformation functions. The behavior of the properties under the transformations is studied and revealed as an important matter to bear in mind. Some examples try to illustrate the proposed framework.

1 Introduction

Similarity and dissimilarity are two widely used concepts in Artificial Intelligence (AI). They appear in many fields like Case Based Reasoning [9], Data Mining [6], Information Retrieval [1], Pattern Matching [14] or Fuzzy Logic [15]. In a general sense, similarity and dissimilarity express a comparison between two elements. Unfortunately, due to this vague definition, there can be different formalizations of similarity and dissimilarity in the literature. Another significative characteristic is the *duality* present between similarity and dissimilarity, as opposite terms but somehow interrelated. This duality extends also to properties, which could be very useful if properly exploited. Thus, every property of a similarity should have a correspondence with one property of a dissimilarity and viceversa.

Several authors have tried to formalize these concepts but main properties are still under discussion [13,2,11]. This lack of basic common theory sometimes leads to incompatible definitions or results focused on an specific kind of similarities or dissimilarities. Duality is commonly ignored as well as some annoying properties (e.g. transitivity) and there are few studies about how transformations of a similarity or dissimilarity can alter their properties.

Therefore, it is fundamental to formalize a definition of both concepts, their main properties, the definition of duality and how they change upon transformations. There is few related work, based on Information Theory [7]. The present contribution extends this formalization of similarity and dissimilarity in two main points. First, with a extended basic formal definition of similarity and dissimilarity and a set of fundamental properties and transformations adaptable to AI fields other than Information Theory. And second, with a study of how this transformations change the properties of the similarities and dissimilarities. The objective of this paper is to point out the significance of setting up a basic framework capable to embrace most of the different theories in order to adapt results from other fields.

The following section introduces the formal definition of similarities and dissimilarities, whereas section 3 and section 4 define the two main transformations between similarities and dissimilarities and their duality. In order to illustrate the approach, in section 5 we introduce some practical examples of similarities, dissimilarities and transformations.

2 Definition

In a general sense, similarity and dissimilarity express the degree of coincidence or divergence between two elements of a reference set. Therefore, it is reasonable to treat them as functions since the objective is to measure or calculate this value between any two elements of the set. In this section, we introduce similarity and dissimilarity functions with their properties as well as further notation. For conciseness, we define similarity and dissimilarity functions at the same time. Note the few but fundamental differences between them.
A similarity function is defined as follows:

Let X be a non-empty set where there is defined an equality relation $\stackrel{X}{=}$. Let s be a function

$$s : X \times X \longrightarrow \mathrm{I}_s \subset \mathbb{R} \tag{1}$$

Assume that s is upper bounded, exhaustive and total. This implies that I_s is upper bounded and also that $\sup_{\mathbb{R}} \mathrm{I}_s$ exists[1].
A dissimilarity function is defined as follows:

Let δ be a function

$$\delta : X \times X \longrightarrow \mathrm{I}_\delta \subset \mathbb{R} \tag{2}$$

Also assume that δ is lower bounded, exhaustive and total. In this case, I_δ is lower bounded and $\inf_{\mathbb{R}} \mathrm{I}_\delta$ exists.

Define $s_{max} \equiv \sup_{\mathbb{R}} \mathrm{I}_s$ and define $\delta_{min} \equiv \inf_{\mathbb{R}} \mathrm{I}_s$. Without loss of generality, consider $s_{max} \geq 0$ and $\delta_{min} \geq 0$. In any other case, a non-negative maximum or minimum can be obtained applying some transformation (e.g. $s + |s_{max}|$). Functions s and δ may be required to satisfy the following axioms, for any $x, y, z \in X$:

Property 1 (Reflexivity)
$s(x, x) = s_{max}$. *This implies* $\sup_{\mathbb{R}} I_s \in I_s$.
$\delta(x, x) = \delta_{min}$. *This implies* $\inf_{\mathbb{R}} I_\delta \in I_\delta$.

[1] In this document we only focus on similarities and dissimilarities whose images are subsets of $\mathbb{R}$. For a more general view see [10].

Property 2 (Strong reflexivity)
$s(x,y) = s_{max} \iff x \overset{X}{=} y$.
$\delta(x,y) = \delta_{min} \iff x \overset{X}{=} y$.

Property 3 (Symmetry)
$s(x,y) = s(y,x)$.
$\delta(x,y) = \delta(y,x)$.

Property 4 (Boundedness)
A similarity s is lower bounded *when $\exists a \in \mathbb{R}$ such that $s(x,y) \geq a$, for all $x, y \in X$. This is equivalent to ask that $\inf_{\mathbb{R}} I_s$ exists.*
A dissimilarity δ is upper bounded *when $\exists a \in \mathbb{R}$ such that $\delta(x,y) \leq a$, for all $x, y \in X$. This is equivalent to ask that $\sup_{\mathbb{R}} I_\delta$ exists.*

Property 5 (Closedness)
Given a lower bounded function s, define $s_{min} \equiv \inf_{\mathbb{R}} I_s$. The property ask for existence of $x, y \in X$ such that $s(x,y) = s_{min}$. This is equivalent to ask that $\inf_{\mathbb{R}} I_s \in I_s$.
Given an upper bounded function δ, define $\delta_{max} \equiv \sup_{\mathbb{R}} I_\delta$. The property ask for existence of $x, y \in X$ such that $\delta(x,y) = \delta_{max}$. This is equivalent to ask that $\sup_{\mathbb{R}} I_\delta \in I_\delta$.

Consider now a function $C : X \longrightarrow 2^X$. Each one of the elements in 2^X will be called a *complement of x*.

Property 6 (Complement) *A lower closed similarity s defined in X has* complement *function C, where $C(x) = \{x' \in X / s(x,x') = s_{min}\}$, if $\forall x, x' \in X$, $|C(x)| = |C(x')| \neq 0$. An upper closed dissimilarity δ defined in X has* complement *function C, where $C(x) = \{x' \in X / \delta(x,x') = \delta_{max}\}$, if $\forall x, x' \in X$, $|C(x)| = |C(x')| \neq 0$.*

Moreover, if s or δ are reflexive, necessarily $x \notin C(x)$. On the other hand, s or δ have unitary complement *function if $\forall x \in X$, $|C(x)| = 1$. In this case, $\forall x \in X$:*

For similarities:	$s(x,y') = s_{max} \iff y' \in C(y), y \in C(x)$
For dissimilarities:	$\delta(x,y') = \delta_{min} \iff y' \in C(y), y \in C(x)$

Let us define a *transitivity operator* in order to introduce transitivity in similarity and dissimilarity functions.

Definition 2.1 (Transitivity operator) *Let I be a non-empty subset of $\mathbb{R}$, and let e be a fixed element of I. A* transitivity operator *is a function $\tau : I \times I \to I$, satisfying the following properties, for all $x, y, z \in I$:*

1. *$\tau(x, e) = x$ (null element)*

2. $y \leq z \Longrightarrow \tau(x,y) \leq \tau(x,z)$ *(non-decreasing monotonicity)*
3. $\tau(x,y) = \tau(y,x)$ *(symmetry)*
4. $\tau(x,\tau(y,z)) = \tau(\tau(x,y),z)$ *(associativity)*

This definition is coincident with the uninorms [4] when $I = [0,1]$. There are two groups of transitivity operators: similarity transitivity operators and dissimilarity transitivity operators. The difference between them is that the former has s_{max} as its null element whereas δ_{min} is the null element of the latter.

Property 7 (Transitivity)
A similarity s defined on X is τ_s-transitive if there is a transitivity operator τ_s such that the following inequality holds:

$$s(x,y) \geq \tau_s(s(x,z), s(z,y)) \quad \forall x,y,z \in X \tag{3}$$

A dissimilarity δ defined on X is τ_δ-transitive if there is a transitivity operator τ_δ such that the following inequality holds:

$$\delta(x,y) \leq \tau_\delta(s(x,z), s(z,y)) \quad \forall x,y,z \in X \tag{4}$$

With all these properties we define similarity and dissimilarity functions: A *similarity* in X is a function s satisfying *strong reflexivity* and *symmetry*. Analogously, a *dissimilarity* in X is a function δ satisfying *strong reflexivity* and *symmetry*. It is easy to prove that *strong reflexivity* implies *reflexivity* and *transitivity* [8].

In this document, the following notation is used: call $\Sigma(X)$ the set of all similarity functions defined over elements of X, call $\Delta(X)$ the set of all dissimilarity functions defined over elements of X. Let us introduce a pair of simple examples of these functions.

Example 2.1 *Let $X = \mathbb{Z}$, and let s a function defined in X:*

$$s(x,y) = 1 - \frac{|x-y|}{|x-y|+1} \tag{5}$$

Where I_s are all the rational numbers on (0,1]. Therefore, $\sup_{\mathbb{Q}} I_s = 1$ and $\inf_{\mathbb{Q}} I_s = 0$. Equality $\overset{X}{=}$ is taken in the usual sense in $\mathbb{Z}$.

This function satisfies reflexivity, symmetry and strong reflexivity axioms. Moreover, it is lower bounded (with $s_{min} = 0$), although it is not lower closed. For this reason, it does not have complement. Also, it fulfills transitivity, expressed as follows:

$$s(x,y) \geq \max\{s(x,z) + s(z,y) - 1, 0\} \tag{6}$$

Example 2.2 *Let X be the set of polygons with n or less vertices, where $n \geq 3$. Consider the relation $x \overset{X}{=} y$ as "x has the same number of vertices than y". Let us define the following dissimilarity*

$$\delta(x,y) = \begin{cases} 0 & \text{if } x \overset{X}{=} y \\ \frac{1}{n - |\#vertices(x) - \#vertices(y)| - 2} & \text{otherwise} \end{cases}$$

Where $I_\delta = \{\frac{1}{n-n_0} \,|\, 2 < n_0 \leq n-1\} \cup \{0\}$. *This dissimilarity has the following properties*[2]*:*

- *Strong Reflexivity and Symmetry. The dissimilarity is the minimum (i.e.* $\delta_{min} = 0$*) only for elements having the same number of vertices. Symmetry is trivial.*
- *Upper boundedness and closedness. Given that, for a fixed number* n *(*$n > 3$*) the dissimilarity between a triangle and a polygon of* n *vertices is* $\delta_{max} = 1$*, this dissimilarity is closed and upper bounded.*
- *Complement. However, the set* $C(x)$ *is empty for those polygons with number of vertices greater than 3 and lower than* n*. For the rest, for all* $x \in X$ *such that* $\#vertices(x) \in \{3, n\}$*:*

$$C(x) \equiv \{x' \in X \quad | \#vertices(x') = |n + 3 - \#vertices(x)|\}$$

Therefore, this dissimilarity does not have complement function.

- *Transitivity. Only fulfilling strong reflexivity and symmetry, this dissimilarity is transitive for the following operator, for all* $x, y, z \in X$*:*

$$\delta(x,y) \leq \frac{\delta(x,z) \cdot \delta(z,y)}{\delta(x,z) + \delta(z,y) - (n-2) \cdot \delta(x,z) \cdot \delta(z,y)}$$

3 Equivalences and Transformations

In the following section we introduce a set of functions that allow us to obtain equivalent similarities or dissimilarities, and to transform a similarity function onto a dissimilarity function and viceversa. All this leads to the concept of duality between similarities and dissimilarities.

3.1 Equivalence Functions

Consider the set of all ordered pairs of elements of X and denote it $X \times X$. For a fixed $s \in \Sigma(X)$, there is a preorder relation in $X \times X$. This preorder is defined as *to belong to a class of equivalence with less or equal similarity value.* This preorder in $X \times X$ depends on s because it is *induced* by s.

Definition 3.1 *Given* X *and* $s \in \Sigma(X)$*, there exists a preorder, denoted by* $\preceq$ *in* $X \times X$*, defined,* $\forall x, y, x', y' \in X$*, as follows.*

$$(x,y) \preceq (x',y') \Longleftrightarrow s(x,y) \leq s(x',y')$$

Analogously, given $\delta \in \Delta(X)$, there exists a preorder in $X \times X$ defined as *to belong to a class of equivalence with less or equal dissimilarity value.* Again, this preorder is induced by δ. Recall that

$$(x,y) \preceq (w,z) \wedge (w,z) \preceq (x,y) \text{ does not imply } x \overset{X}{=} w \wedge y \overset{X}{=} z.$$

[2] This dissimilarity does not reflect the usual concept of similarity in geometry. It is only for illustrative purposes.

Using these concepts, the definition of equivalence between similarities and equivalence between dissimilarities is introduced.

Definition 3.2 (Equivalent similarities/disimilarities) *Two similarities (dissimilarities) with the same reference set X are* equivalent *if they induce the same preorder in $X \times X$.*

Note that the equivalence between similarities or between dissimilarities is an equivalence relation.

The properties of similarities and dissimilarities are kept under equivalence, except boundedness.

Proposition 3.1 *Given two equivalent similarities $s_1, s_2 \in \Sigma(X)$ or two equivalent dissimilarities $\delta_1, \delta_2 \in \Delta(X)$,*

- s_1 *(*δ_1*) is* reflexive *only if* s_2 *(*δ_2*) is* reflexive*.*
- s_1 *(*δ_1*) is* strong reflexive *only if* s_2 *(*δ_2*) is* strong reflexive*.*
- s_1 *(*δ_1*) is* symmetric *only if* s_2 *(*δ_2*) is* symmetric*.*
- s_1 *(*δ_1*) is* lower closed *(*upper closed*) only if* s_2 *(*δ_2*) is* lower closed *(*upper closed*).*
- s_1 *(*δ_1*) has* (unitary) complement *only if* s_2 *(*δ_2*) has (*unitary*)* complement*, with the same complement function.*
- s_1 *(*δ_1*) is* transitive *only if* s_2 *(*δ_2*) is* transitive*.*

Proof. The proof can be found in [8].

Recalling Definition 3.1 it is easy to see that only the *monotonically increasing* and *invertible* functions keep the preorder induced. In this sense, we define an equivalence function that allows us to get equivalent similarities or dissimilarities.

Definition 3.3 (Equivalence function) *Let s_1 be a similarity and δ_1 be a dissimilarity. An* equivalence function *is a monotonically increasing and invertible function $\breve{f}$ such that $s_2 = \breve{f} \circ s_1$ is a similarity on $X \times X$ equivalent to s_1. Analogously, $\delta_2 = \breve{f} \circ \delta_1$ is a dissimilarity on $X \times X$ equivalent to δ_1.*

Furthermore, if τ_1 is the transitivity operator of s_1 (or δ_1), then the transitivity operator of s_2 (or δ_2) is given by the following expression:

$$\tau_2(a, b) = \breve{f}(\tau_2(\breve{f}^{-1}(a), \breve{f}^{-1}(b))) \tag{7}$$

Therefore, any composition of an equivalence function and a similarity (or dissimilarity) function is another similarity (or dissimilarity) function, which is also equivalent.

Let us introduce a special set of similarities and dissimilarities. Denote $\Sigma^{\star}(X)$ the set of similarities defined on X with codomain on [0,1] and denote $\Delta^{\star}(X)$ the set of similarities defined on X with codomain on [0,1]. Denote $\breve{f}^{\star}$ the subset of equivalence functions such that its domain or its codomain is [0,1]. Using these functions, we have a way to get equivalent similarities or dissimilarities on

$\Sigma^\star(X)$ using similarities or dissimilarities on $\Sigma(X)$ and viceversa. Thus, defining properties of similarities and dissimilarities on $\Sigma^\star(X)$ and $\Delta^\star(X)$ is the same that defining them on $\Sigma(X)$ and $\Delta(X)$. The important reason to do this is that there is a lot of previous work on functions in [0,1] that we can adapt to similarities and dissimilarities [12,5].

3.2 Transformation Functions

Equivalence functions allow us get new similarities from other similarities or new dissimilarities from other dissimilarities, but not to change between the former and the latter. Consider the previous sets $\Sigma^\star$ and $\Delta^\star$, with similarities and dissimilarities defined on [0,1]. Next definitions introduce the concept of *transformation function* between similarities and dissimilarities.

Definition 3.4 (Transformation function on [0,1]) *A* transformation function $\hat{n}$ *is a* decreasing *bijection on [0,1]. This implies the following:*

- $\hat{n}(0) = 1$ *and* $\hat{n}(1) = 0$*. These are called* limit conditions.
- $\hat{n}$ *is continuous on [0,1].*
- $\hat{n}$ *has inverse on [0,1].*

A transformation function is involutive *if* $\hat{n}^{-1} = \hat{n}$*.*

Note that this definition is restricted to similarities and dissimilarities in $\Sigma^\star(X)$ and $\Delta^\star(X)$. However, using that both $\breve{f}^\star$ and $\hat{n}$ are bijections an analog transformation function between elements of $\Sigma(X)$ and $\Delta(X)$ is the composition of two or more functions in the following way:

Definition 3.5 (Transformation function) *A* transformation function *(denoted $\hat{f}$) is defined as a composition of two equivalence functions and a transformation function on* $[0,1]$*.*

$$\hat{f} = \breve{f}_1^\star \circ \hat{n} \circ {\breve{f}_2^\star}^{-1}$$

Where $\hat{n}$ is a transformation function on $[0,1]$, $\breve{f}_1^\star$ gets an equivalent similarity or dissimilarity on $\Sigma(X)$ or $\Delta(X)$ whereas $\breve{f}_2^\star$ gets an equivalent similarity or dissimilarity on $\Sigma^\star(X)$ or $\Delta^\star(X)$.

4 Duality

As it has been seen along this paper, similarity and dissimilarity are two interrelated concepts. In fuzzy theory [5], t-norms and t-conorms are dual with respect to the fuzzy complement. In the same sense, similarity and dissimilarity functions are *dual* with respect to some transformation function.

Definition 4.1 (Duality) *Given $s \in \Sigma(X)$, $\delta \in \Delta(X)$ and a transformation function $\hat{f}$. We say that s, δ are* dual *by means of $\hat{f}$ if*

$$\delta = \hat{f} \circ s \tag{8}$$

or the equivalent form

$$s = \hat{f}^{-1} \circ \delta \tag{9}$$

This duality is expressed by the triple $< s, \delta, \hat{f} >$.

Theorem 1. *Given a dual triple $< s, \delta, \hat{f} >$, it is true that*

- *δ is strongly reflexive only if s is strongly reflexive.*
- *δ is closed only if s is closed.*
- *δ has (unitary) complement only if s has (unitary) complement.*
- *δ is τ_δ-transitive only if s is τ_s-transitive, where τ_δ can be defined by means of τ_s*

$$\tau_\delta(x, y) = \hat{f}(\tau_s(\hat{f}^{-1}(x), \hat{f}^{-1}(y))) \tag{10}$$

Proof. *The proof can be found in [8].*

Via this duality we can define properties or theorems on similarities which are immediately applicable also to dissimilarities, or viceversa. A general view of all the functions and sets appeared so far is represented in figure 4.1.

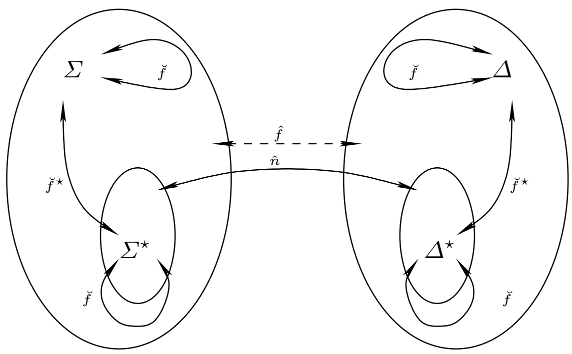

Fig. 4.1. Graphical representation of equivalence ($\check{f}$) and transformation ($\hat{f}$) functions in Σ and Δ.

5 Examples

Let us introduce some illustrative examples. Two equivalent similarities or two equivalent dissimilarities fulfill almost the same properties. The exception is the bound property which depends on the equivalence function chosen. Although transitivity is always kept, the operator that holds the property is also dependent on the equivalence function applied.

Example 5.1 *A metric dissimilarity* $\delta_1 \in \Delta(X)$ *is symmetric, bounded and fulfills the triangle inequality. If we apply the function* $\hat{f}(z) = 1 - z$ *to* δ_2 *we get a similarity* s_1 τ_{s_1}*-transitive (i.e.* $\tau_{s_1}(a, b) = a + b - 1$*). However, if we apply the function* $\breve{f}(z) = z^2$ *to* δ_1 *we get a dissimilarity but not a metric.*

Besides, if we apply $\breve{f}$ *to* s_1*, we get a similarity with a transitivity stronger than* s_1*. Since*

$$(a + b - 1)^2 \geq (a)^2 + (b)^2 - 1$$

then $\tau_{s_2} \sqsupset \tau_{s_1}$*. Simplifying we get:*

$$ab + 1 \geq a + b$$

This is always true if $a, b \in [0, 1]$*. Denote* s_2 *to this new similarity.*

Now consider the following dissimilarity:

$$\delta_2 = 1 - s_1$$

Which is the relation between δ_1 *and* δ_2*? Obviously* $\delta_2 = \breve{f}_2 \circ \delta_1 = \delta_1(2 - \delta_1)$*. However, if* $\tau_{s_2} \sqsupset \tau_{s_1}$*, which is the relation between* τ_{δ_1} *and* τ_{δ_2}*? Using that* $\breve{f}_2$ *is subadditive we state that* $\tau_{\delta_2} \sqsupset \tau_{\delta_1}$*, this is* τ_{δ_2} *is* stronger *than* τ_{δ_1}*. This means that* δ_2 *is also a metric dissimilarity.*

In the next table are collected the similarities and dissimilarities described here and their respective transitivity operators.

$$\begin{aligned}
\delta_1(x, y) &= |x - y| \\
\tau_{\delta_1}(a, b) &= \min\{a + b, 1\} \\
s_1(x, y) &= 1 - |x - y| \\
\tau_{s_1}(a, b) &= \max\{a + b - 1, 0\} \\
s_2(x, y) &= 1 + (x - y)^2 - 2|x - y| \\
\tau_{s_2}(a, b) &= (\sqrt{a} + \sqrt{b} - 1)^2 \\
\delta_2(x, y) &= 2|x - y| - (x - y)^2 \\
\tau_{\delta_2}(a, b) &= 1 - (\sqrt{1 - a} + \sqrt{1 - b} - 1)^2
\end{aligned}$$

Figure 5.2 illustrates all the process.

This behavior of transitivity and how it is affected by the equivalence and transformation functions is an interesting field of study. Some works have been previously done in this sense using metrics (preservation of metric properties) [3] and uninorms (gradation of transitivity operators). In this case, similarity and dissimilarity unify both fields in a general view of preservation of transitivity using equivalence functions. This transformation of transitivity can be used, for example, to get a metric dissimilarity from a non-metric one. In the following example we compare the structure of two trees with a non-metric dissimilarity, applying an equivalent function we get an equivalent and metric dissimilarity function.

Example 5.2 *Consider a dissimilarity function between binary trees. It does not measure differences between nodes but the structure of the tree. First, let us*

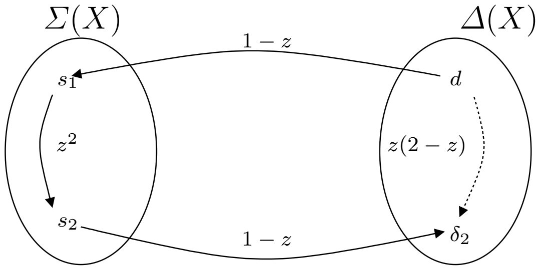

Fig. 5.2. Transformation corresponding to Example 5.1

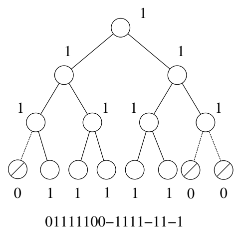

Fig. 5.3. A simple coding of binary trees. The reason of going bottom-up is to have the less significative digits close to the root of the tree. The choice of making the left nodes more significant than the right ones is arbitrary. $\oslash$ represents the empty tree.

describe a tree coding function D. It assigns a unique value for each tree, this value is first coded as a binary number and later interpreted as a natural number, and contains $2^h - 1$ bits, where h is the height of the tree. The value is calculated such that the most significant bit corresponds to the leftmost and bottommost tree node (see figure 5.3). Note that D is not a bijection, since there are numbers that do not code any valid binary tree.

Consider now the following dissimilarity function, where A and B are binary trees; $\oslash$ represents the empty tree and its value is 0.

$$\delta(A,B) = \begin{cases} \max(\frac{D(A)}{D(B)}, \frac{D(B)}{D(A)}) & \text{if } A \neq \oslash \text{ and } B \neq \oslash \\ 1 & \text{if } A = \oslash \text{ and } B = \oslash \\ D(A) & \text{if } A \neq \oslash \text{ and } B = \oslash \\ D(B) & \text{if } A = \oslash \text{ and } B \neq \oslash \end{cases}$$

This a strong reflexive, symmetric, unbounded dissimilarity defined on $[1, +\infty)$ with $\delta_{min} = 1$. If we impose a limit to the height of the trees (call it H), then δ is upper bounded and closed with $\delta_{max} = \sum_{i=0}^{2^H-1} 2^i$.

It is also transitive with the product operator (which is a dissimilarity transitivity operator on $[1, +\infty)$[3]). The following inequality is always true for any

[3] The product with elements on $[1, +\infty)$ is an associative, symmetric, monotonically non-decreasing function with 1 as null element.

three trees A, B, C.

$$\delta(A,B) \leq \delta(A,C) \cdot \delta(C,B)$$

If A, B *and* C *are not the empty tree, substituting in the previous expression and operating with* max *and the product we get:*

$$\max\left(\frac{D(A)}{D(B)}, \frac{D(B)}{D(A)}\right) \leq \max\left(\frac{D(A)}{D(B)}, \frac{D(C)^2}{D(A)\cdot D(B)}, \quad \frac{D(A)\cdot D(B)}{D(C)^2}, \frac{D(B)}{D(A)}\right)$$

This is trivially true.

If $A = \oslash$*, then the inequality is* $D(B) \leq \max\left(D(B), \frac{D(C)^2}{D(B)}\right)$ *which is trivially true. It is also true for* $B = \oslash$ *or* $C = \oslash$*. This ends the proof.*

If we apply the following equivalence function $\breve{f}(z) = \log z$ *to* δ *we get an* equivalent *dissimilarity function* $\delta' = \breve{f} \circ \delta$*. By Proposition 3.1 the properties of* δ *are kept in* δ'*. However, the transitivity operator is changed using Expression 7, as* $\tau_{\delta'}(a,b) = a + b$*. Therefore, we have a metric dissimilarity function over trees and fully equivalent to the initial choice of* δ*.*

6 Conclusions

The main goal of this paper has not been to set up a standard definition of similarity and dissimilarity, but to establish some grounds on the definition of these widely used concepts. Also we have tried to set up fundamental transformations in order to keep the basic properties. However, a deeper study has to be done about the effects of transformations, specially in transitivity (e.g. which transformations do keep the triangle inequality) and more complex matters, like aggregation of different measures into a global one. There is a lot of work done in Fuzzy Theory or Information Retrieval that can be adapted to this framework and, therefore, adapted in turn to other AI related fields. Due to the many fields of application these concepts are involved with, the study of their properties can lead to better understanding of similarity and dissimilarity measures in many areas of AI. We think this paper can be useful as a formal basis to a wider and deeper study of similarity and dissimilarity theory.

Acknowledgements

This document has been granted by the Interministerial Council of Science and Technology (CICYT) of Spain, within the project DPI2002-03225 and by the Spanish Ministry of Education, Culture and Sport (MECD) with the grant FPU AP2001-3735.

References

1. R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*, ACM Press New York, 1999.

2. M. De Cock and E. Kerre, 'On (un)suitable relations to model approximate equality', *Fuzzy Sets And Systems*, **133**, 137–153, (2003).
3. Michiel Hagedoorn and Remco C. Veltkamp, 'Reliable and efficient pattern matching using an affine invariant metric', *International Journal of Computer Vision*, **2/3**(31), 203–225, (1999).
4. E. P. Klement, 'Some mathematical aspects on fuzzy sets: Triangular norms, fuzzy logics and generalized measures', *Fuzzy Sets And Systems*, **90**, 133–140, (1997).
5. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, chapter 3, Pearson Education POD, 1995.
6. T. Li, Z. Shenghuo, and M. Ogihara, 'A new distributed data mining model based on similarity', in *ACM SAC Data Mining Track*, Melbourne, Florida, USA, (March 2003).
7. D. Lin, 'An information-theoretic definition of similarity', in *Proceedings of International Conference on Machine Learning*, Madison, Wisconsin, (July 1998).
8. J. Orozco, 'Similarity and dissimilarity concepts in machine learning', Technical Report LSI-04-9-R, Universitat Politécnica de Catalunya, Spain, (2004).
9. H. Osborne and D. Bridge, 'Models of similarity for case-based reasoning', in *Procs. of the Interdisciplinary Workshop on Similarity and Categorisation*, eds., E.Cambouropolos M.Ramscar, U.Hahn and H.Pain, pp. 173–179, (1997).
10. H. Osborne and D. Bridge, 'Similarity metrics: A formal unification of cardinal and non-cardinal similarity measures', in *ICCBR*, pp. 235–244, (1997).
11. S. Santini and R. Jain, 'Similarity measures', in *IEEE Transactions On Pattern Analysis and Machine Intelligence*, (1999).
12. B. Schweizer and A. Sklar, *Probabilistic Metric Spaces*, North-Holland, Amsterdam, 1983.
13. A. Tversky, *Features Of Similarity*, volume 84(4), 327–352, Psycological Review, July 1977.
14. R. C. Veltkamp and M. Hagedoorn, 'Shape similarity measures, properties and constructions', in *Advances in Visual Information Systems, 4th International Conference, VISUAL 2000*, volume 1929 of *Lecture Notes in Computer Science*. Springer, (2000).
15. J. Williams and N. Steele, 'Difference, distance and similarity as a basis for fuzzy decision support based on prototypical decision classes', *Fuzzy Sets and Systems*, **131**, 35–46, (2002).