Towards a Conscious Choice of a Fuzzy Similarity Measure: A Qualitative Point of View

Bernadette Bouchon-Meunier[1], Giulianella Coletti[2], Marie-Jeanne Lesot[1], and Maria Rifqi[1]

[1] Université Pierre et Marie Curie - Paris 6, CNRS UMR 7606, LIP6, 104 avenue du Président Kennedy, F-75016 Paris, France
{Bernadette.Bouchon-Meunier,Marie-Jeanne.Lesot,Maria.Rifqi}@lip6.fr
[2] Dipartimento Matematica e Informatica, Università di Perugia, Via Vanvitelli 1, 06123 Perugia, Italy
coletti@dipmat.unipg.it

Abstract. In this paper, we propose to study similarity measures among fuzzy subsets from the point of view of the ranking relation they induce on object pairs. Using a classic method in measurement theory, introduced by Tversky, we establish necessary and sufficient conditions for the existence of a class of numerical similarity measures, to represent a given ordering relation, depending on the axioms this relation satisfies.

Keywords: Fuzzy similarity, comparison measure, ordering relation, representability, weak independence conditions.

1 Introduction

Similarity is a key concept in artificial intelligence [12] and similarity measures have been extensively studied (see [7,14,2,1], see also the surveys [5,8]). The choice of an appropriate measure when facing a particular problem to solve is a central issue. Now, due to the subjective characteristic of similarity as used by human beings, it is more intuitive to compare measures depending on the order they induce rather than the numerical values they take. Therefore, trying to get closer to the human reasoning, we propose to consider an ordinal view on similarity measures. To that aim, we follow the approach proposed by Tversky [14] and applied later in [3], in the framework of measurement theory: it starts from a *comparative similarity* $\preceq$, defined as a binary relation on object pairs, and studies the conditions under which $\preceq$ can be represented by a numerical similarity measure. It establishes representation theorems that state necessary and sufficient conditions under which a given comparative similarity is represented by a specific form of numerical similarity measures.

Previous works [14,3] considered the crisp case of presence/ absence data, we focus in this paper on fuzzy data: for any object, the presence of an attribute is not binary but measured by a membership degree in [0,1]. Considering such fuzzy data raises several difficulties due to the associated softness and change continuity: in the crisp case, for any object pair and any attribute, only four

E. Hüllermeier, R. Kruse, and F. Hoffmann (Eds.): IPMU 2010, LNAI 6178, pp. 1–10, 2010.

configurations can occur, namely whether the attribute is present in both objects, absent from both, or present in one object but not the other one. Moreover, only two kinds of modifications can occur, changing an attribute presence to an absence or reciprocally. In the fuzzy case, all modifications are continuous, and it is not possible to identify a finite set of distinct configurations. Thus matching the ordinal view of similarity with the numerical one requires the definition of new properties and axioms to characterise the possible behaviors of comparative similarity. Furthermore, in the fuzzy framework, similarity measures cannot be reduced to their general form: they also depend on the choice of a t-norm and a complementation operator, to define the membership degrees to the intersection and the complement of fuzzy sets respectively. Indeed, as illustrated in Section 2, changing the t-norm can lead to very different comparative relations for a given similarity measure form and a given fuzzy measure. This implies that the axioms we introduce to characterise comparative similarities depend on the t-norm choice. We consider the three most common t-norms (min, product and Łukasiewicz t-norm) and characterise the comparative similarities representable by (or agreeing with) a class of similarity measures containing as particular elements Jaccard, Dice, Sorensen, Anderberg and Sokal-Sneath measures.

The paper is organised as follows: in Section 2, we recall the definitions of similarity measure representation and equivalence and we introduce basic axioms, in particular those expressing constraints in terms of attribute uniformity and monotonicity. Section 3 presents the considered independence axiom that is required to establish, in Section 4, the representation theorem.

2 From Numerical Similarity to Comparative Similarity

In this section, after introducing the notations used throughout the paper, we discuss the classic definition of equivalence between numerical similarity measures and establish basic axioms satisfied by comparative similarities induced from given classes of numerical similarities, following the ideas of Tversky to study similarity using the framework of measurement theory [14].

2.1 Preliminaries

We consider that each object is described by p attributes, i.e. by the set of characteristics from the predefined list $\mathcal{A}$, which can be present with different degrees of membership: any object is a fuzzy subset of $\mathcal{A}$. The data set is noted $\mathcal{X} = [0,1]^p$: any $X \in \mathcal{X}$ is written $X = (x_1, ..., x_p)$, $x_i \in [0,1]$, and associated with $s_X = \{i : x_i > 0\}$; $\underline{0}$ denotes the object with $s_X = \emptyset$. We consider a t-norm $\top$ and its dual t-conorm $\bot$ and the complement $X^c = 1 - X$. We define, as usual, $X \cap Y = X \top Y$, $X \setminus Y = X \top Y^c$ and $Y \setminus X = Y \top X^c$. We say that X^* is a *strong* $\top$*-complement* of X if $s_{X \cap X^*} = \emptyset$, when the intersection is ruled by the t-norm $\top$.

For any $0 \leq \delta \leq x_i$ and $0 \leq \eta \leq 1 - x_i$ we denote by $x_i^{-\delta} = x_i - \delta$, by $x_i^{\eta} = x_i + \eta$, and by $X_k^{-\delta} = \{x_1, ..., x_k^{-\delta}, ..., x_p\}$. Lastly, given η such that $0 \leq \eta \leq \min_i(1 - x_i)$, we note $X^{\eta} = \{x_1^{\eta}, ..., x_p^{\eta}\}$.

We indicate by m a measure of fuzzy sets, only depending on the values of the membership and not on the values of the considered domain, for instance $m(X) = \Sigma_i x_i$ ($m(X) = \int_{S_X} X(t)dt$ in the infinite case). Given $X, Y \in \mathcal{X}$, we note $\mathbf{x} = m(X \cap Y), \mathbf{y}^+ = m(X \setminus Y), \mathbf{y}^- = m(Y \setminus X)$ and $\mathbf{y} = \mathbf{y}^+ + \mathbf{y}^-$.

We then consider a *comparative similarity*, defined as a binary relation $\preceq$ on $\mathcal{X}^2$, with the following meaning: for $X, Y, X', Y' \in \mathcal{X}$, $(X, Y) \preceq (X', Y')$ means that X is similar to Y no more than X' is similar to Y'.

The relations $\sim$ and $\prec$ are then induced by $\preceq$ as: $(X, Y) \sim (X', Y')$ if $(X, Y) \preceq (X', Y')$ and $(X', Y') \preceq (X, Y)$, meaning that X is similar to Y as X' is similar to Y'. Lastly $(X, Y) \prec (X', Y')$ if $(X, Y) \preceq (X', Y')$ holds, but not $(X', Y') \preceq (X, Y)$. If $\preceq$ is complete, then $\sim$ and $\prec$ are the symmetrical and the asymmetrical parts of $\preceq$ respectively.

We now introduce the notion of representability for such a comparative similarity by a numerical similarity measure:

Definition 1. *Given a comparative similarity $\preceq$, a similarity measure $S : \mathcal{X}^2 \to \mathbb{R}$ represents $\preceq$ if and only if $\forall (X, Y), (X', Y') \in \mathcal{X}^2$*

$$(X, Y) \preceq (X', Y') \Leftrightarrow S(X, Y) \leq S(X', Y')$$

It is important to notice that in a fuzzy context, if the similarity measure involves the intersection, union or difference of the fuzzy sets whose similarity is studied, then we need to consider also the particular t-norm and t-conorm we choose: the induced comparative similarity can change for different choices of t-norm and t-conorm, as the following example shows: let us consider $\mathcal{X} = [0, 1]^4$, $X = (0, 0, 2/10,\ 3/10)$, $Y = (0, 0, 4/10, 1/10)$, $U = (0, 7/10, 0, 2/10)$, $V = (0, 0, 3/10, 9/10)$, $W = (0, 0, 1/10, 3/10)$, $Z = (0, 2/10, 4/10, 0)$. As a fuzzy measure, we choose $m(X) = \Sigma_i x_i$ and as a similarity measure

$$S = S_\rho(X, Y) = \frac{\mathbf{x}}{\mathbf{x} + \rho \mathbf{y}} \tag{1}$$

with $\rho = 1$ and we indicate by $S_\top$ the measure S when we choose the t-norm $\top$.

If we adopt min as a t-norm, then we obtain: $S_m(X, Y) = 3/13 > S_m(U, V) = 2/21 > S_m(Z, W) = 1/11$, and so $(Z, W) \prec (U, V) \prec (X, Y)$;

If we adopt Łukasiewicz t-norm, we obtain: $S_L(U, V) = 1/18 > S_L(X, Y) = S_L(Z, W) = 0$, and so $(X, Y) \sim (Z, W) \prec (U, V)$;

If we adopt the product as t-norm, we obtain: $S_p(X, Y) = 11/89 > S_p(U, V) = 3/32 > S_p(Z, W) = 1/24$, and so $(Z, W) \prec (X, Y) \prec (U, V)$.

2.2 Similarity Measure Equivalence

Any similarity measure on $\mathcal{X}^2$ induces a comparative similarity $\preceq$, defined as follows: $(X, Y) \prec (X', Y')$ if $S(X, Y) < S(X', Y')$ and $(X, Y) \sim (X', Y')$ if $S(X, Y) = S(X', Y')$.

Now the same ordering relation is induced by any similarity measure that can be expressed as an increasing transformation of S: any similarity measure

$S' = \varphi(S)$, with $\varphi : \mathbb{R} \to \mathbb{R}$ strictly increasing is also a representation of $\preceq$. Moreover, no other measure represents $\preceq$, as shown in [9,10]. Thus, from a comparative point of view, all functions $\varphi(S)$ are indistinguishable. Formally speaking, the relation r defined on the set of similarity measures as SrS' if and only if S and S' induce the same comparative similarity on $\mathcal{X}$ is an equivalence relation. An equivalent formulation of this concept is given in [9,10].

The similarity measures defined by Equation 1, with $\rho > 0$, are all equivalent, since each of them is an increasing transformation of any other. In particular, the Jaccard ($\rho = 1$), Dice ($\rho = 1/2$), Sorensen ($\rho = 1/4$), Anderberg ($\rho = 1/8$) and Sokal and Sneath ($\rho = 2$) measures are equivalent.

The same class also contains the function $S(X,Y) = \log(\mathbf{x}) - \log(\mathbf{y})$, which is of the kind proposed by Tversky [14]: S is an increasing transformation of $S'(X,Y) = \mathbf{x}/\mathbf{y}$ which is an increasing transformation of S_1.

It is to be noted that the function $S(X,Y) = \alpha \log(\mathbf{x}) - \beta \log(\mathbf{y})$ for $\alpha, \beta > 0$ is not in the same class, but it is equivalent to all measures

$$S^*_\rho(X,Y) = \frac{\mathbf{x}^\alpha}{\mathbf{x}^\alpha + \rho \mathbf{y}^\beta}$$

Obviously all these considerations hold for any particular choice of t-norm.

2.3 Basic Axioms

We are now interested in a different classification of similarity measures: instead of considering the measures that induce the same order, we consider the measures that induce orders satisfying the same class of axioms. In this section, we consider axioms that lead to preliminary results regarding relations between similarity measures and comparative similarity.

The first two axioms we introduce describe basic properties a binary relation has to satisfy to define a comparative similarity: the first one only states the relation must be a weak order and the uncountable set $\mathcal{X}$ has a countable subset thoroughly interspersed.

Axiom S1 [weak order]
The relation $\preceq$ defined on $\mathcal{X}$ is a weak order, i.e it is complete, reflexive and transitive.
There exists a countable set $\mathcal{Y} \subseteq \mathcal{X}$ which is dense in $\mathcal{X}$ with respect to $\preceq$.

We recall that any comparative structure representable by a real function satisfies S1 (see for instance [6], Theorem 2). We note that if we require that the membership values are rational (and so in particular in the crisp case), then the second part of S1 is automatically satisfied, since in this case $\mathcal{X}$ itself is countable.

The second axiom expresses boundary conditions.

Axiom S2 [boundary conditions]
$\forall X, X', Y, Y' \in \mathcal{X}$, with $s_{X \cap X'} = s_{Y \cap Y'} = \emptyset$
$(X, X') \sim (Y, Y') \preceq (X, Y) \preceq (X, X) \sim (Y, Y)$

Axiom S2 requires that any X differs from any Y no more than from its strong $\top$-complements, and not less than itself and it imposes that X is similar to itself as Y is to itself.

It is obvious that S2 strictly depends on the chosen t-norm: in particular for min and product the class of minimal elements coincides with the class of elements such that $s_X \cap s_{X'} = s_Y \cap s_{Y'} = \emptyset$. For Łukasiewicz t-norm it is larger, since we can obtain 0 also starting from two positive values. The axiom moreover implies that, if the membership of the intersection of two object descriptions is null, then it is indifferent whether the attributes are absent of both objects or present in one of them, with any degree of membership. In particular, if we use Łukasiewicz t-norm, two objects having all attributes in common, but with the sum of involved degrees less than 1, are as similar as two objects having each attribute completely present in one of them and completely absent from the other one. This makes a major difference with the crisp case.

The third axiom imposes a symmetry condition.

Axiom S3 [symmetry]
$\forall X, Y \in \mathcal{X}$, $(X, Y) \sim (Y, X)$

These properties lead to the following two definitions:

Definition 2. *A binary relation $\preceq$ on $\mathcal{X}^2$ is a* comparative fuzzy similarity *if and only if it satisfies axioms S1 and S2.*

Definition 3. *A comparative fuzzy similarity is* symmetric *if and only if it satisfies axiom S3.*

The next axiom, named attribute uniformity, examines the conditions under which the attributes can be considered as having the same role with respect to the comparative similarity, i.e. the conditions under which a modification in one attribute is equivalent to the modification of another attribute. In the crisp case, this axiom only has to consider two kinds of modifications, changing an attribute presence to an absence or reciprocally. Moreover, it only has to examine the four categories the attributes belong to (whether they are present in both objects, absent from both, or present in one object but not the other one). In the fuzzy case, all modifications are continuous and their effects cannot be simply expressed as a transition between such categories. The attribute uniformity axiom depends on the considered t-norm that determines the admissible modifications to be considered. Thus the axiom takes three formulations detailed below.

Axiom $S4_{min}$ [attribute uniformity]
$\forall\, h, k \in \{1, ..., p\}$, such that $x_h \geq y_h$ and $x_k \geq y_k$, and $x_h \geq 1 - y_h$ and $x_k \geq 1 - y_k$
and for all real numbers $\varepsilon, \eta, \vartheta, \gamma$, with
$0 \leq \varepsilon \leq \min_{i=h,k}\{(x_i - y_i), (y_i + x_i - 1)\}$; $0 \leq \eta \leq \min_{i=h,k}\{(1 - x_i)\}$,
$0 \leq \vartheta \leq \min_{i=h,k}\{(y_i + x_i - 1)\}$; $0 \leq \gamma \leq \min_{i=h,k}\{(x_i - y_i)\}$.
one has:

$$(X_k^{-\varepsilon}, Y) \sim (X_h^{-\varepsilon}, Y), \qquad (X_k^{\eta}, Y) \sim (X_h^{\eta}, Y)$$
$$(X, Y_k^{-\vartheta}) \sim (X, Y_h^{-\vartheta}), \qquad (X, Y_k^{\gamma}) \sim (X, Y_h^{\gamma})$$

The same condition holds for $x_h \leq 1 - y_h$ and $x_k \leq 1 - y_k$ and $0 \leq \varepsilon \leq \min_{i=h,k}\{(x_i - y_i)\}$; $0 \leq \eta \leq \min_{i=h,k}\{(1 - y_i - x_i)\}$ $0 \leq \vartheta \leq \min_{i=h,k}\{(y_i)\}$; $0 \leq \gamma \leq \min_{i=h,k}\{(x_i - y_i), (1 - y_i - x_i)\}$.

Moreover for $\alpha \leq (x_k - y_k)$ if $x_h \geq 1 - y_h$ and $x_k \geq 1 - y_k$, and $\alpha \leq \min\{(x_k - y_k), (1 - y_k - x_k)\}$ if $x_h \leq 1 - y_h$ and $x_k \leq 1 - y_k$, one has:

$$(X, Y) \sim (X, (Y_h^{-\alpha})_k^{\alpha}).$$

Axiom $S4_p$ [attribute uniformity] $\forall h, k \in \{1, ..., p\}$, such that $y_k = y_h$, and for all real numbers ε, η, with $0 \leq \varepsilon \leq \min_{i=h,k}\{x_i\}$, $0 \leq \eta \leq \min_{i=h,k}\{1 - x_i\}$, one has:

$$(X_k^{-\varepsilon}, Y) \sim (X_h^{-\varepsilon}, Y); \quad (X_k^{\eta}, Y) \sim (X_h^{\eta}, Y)$$

A symmetric condition holds if $x_k = x_h$.
Moreover $((X_h^{-\varepsilon})_k^{\varepsilon}, Y) \sim (X, Y)$.

Axiom $S4_L$ [attribute uniformity] $\forall h, k \in \{1, ..., p\}$, such that $y_i < x_i$, $x_i \geq 1 - y_i$ $i = h, k$, and for all real numbers ε, γ, with $0 \leq \varepsilon \leq \min_{i=h,k}\{(x_i - y_i), (y_i + x_i - 1)\}$; $0 \leq \gamma \leq \min_{i=h,k}\{(x_i - y_i)\}$. one has

$$(X_k^{-\varepsilon}, Y) \sim (X_h^{-\varepsilon}, Y); \quad (X, Y_k^{\gamma}) \sim (X, Y_h^{\gamma}).$$

The same condition holds for $x_i \leq 1 - y_i$ and $x_i \leq 1 - y_i$, $i = h, k$, and $0 \leq \varepsilon \leq \min_{i=h,k}\{(x_i - y_i)\}$ $0 \leq \gamma \leq \min_{i=h,k}\{(x_i - y_i), (1 - y_i - x_i)\}$.

We can prove the following Proposition:

Proposition 1. *Let us consider on $\mathcal{X}^2$ the t-norm $\top$ ($\top = min, p, L$) and the comparative fuzzy similarity $\preceq$ satisfying Axiom $S4_\top$. If $m : \mathcal{X}^2 \to \mathbb{R}$ is a fuzzy measure such that $m(X) = \varphi(\sum_i x_i)$ with φ an increasing real function, then the following condition holds:*

if for (X, Y, Z) one has $\mathbf{x} = \mathbf{z}$, $\mathbf{x}^- = \mathbf{z}^-$ and $\mathbf{x}^+ = \mathbf{z}^+$ then $(X, Y) \sim (Z, Y)$

Proof: let us consider the case $\top = \min$. First we note that, starting from (X, Y), we can obtain (Z, Y) with a sequence of steps as those considered in $S4_{min}$. For each of them, we have equivalent pairs or pairs with the same values for $\mathbf{x}$, $\mathbf{x}^-$ and $\mathbf{x}^+$. We prove this in the case $(X_k^{-\varepsilon}, Y)$ and $(X_h^{-\varepsilon}, Y)$, with $x_h \geq 1 - y_h$ and $x_k \geq 1 - y_k$: we have in fact: $m(X_k^{-\varepsilon} \cap Y) = m(X_h^{-\varepsilon} \cap Y) = m(X \cap Y)$; $m(X_k^{-\varepsilon} \setminus Y) = m(X_h^{-\varepsilon} \setminus Y) = m(X \setminus Y)$; $m(Y \setminus X_k^{-\varepsilon}) = m(Y \setminus X_h^{-\varepsilon}) = m(X \setminus Y) + \varepsilon$. Similar proofs can be obtained for the other cases.

It must be underlined also that this axiom is satisfied by any comparative similarity representable by a similarity measure, only depending on $\mathbf{x} = m(X \cap Y), \mathbf{y}^+ = m(X \setminus Y), \mathbf{y}^- = m(Y \setminus X)$ and $\mathbf{y} = \mathbf{y}^+ + \mathbf{y}^-$, where m is a measure of fuzzy sets, depending on the values of the membership and not on the values of the considered domain. Reciprocally, any comparative similarity satisfying S4 can be represented by a function depending only on $\mathbf{x}, \mathbf{y}^+, \mathbf{y}^-$.

2.4 Monotonicity Axioms

The following three axioms of monotonicity govern the comparative similarity among pairs differing in a different degree of belonging of only one attribute, successively for the three considered t-norms.

Axiom $S5_{min}$ **[monotonicity]** $\forall X, Y \in \mathcal{X},\ X \neq Y$
$\forall\, k, h, j \in s_X \cap s_Y$, such that $y_r > x_r\,, (r = k, h)\,,\, x_j > y_j\,,\, 1 - x_s < y_s\,, (s = k, j)\,,\, 1 - x_h > y_h$ (so $y_k, x_j > 1/2, x_h < 1/2$).
For any $0 < \varepsilon < \min\{x_h, x_k, y_j, 2(y_h + x_h - 1), (y_k + x_k - 1)/2, (x_j + y_j - 1)\}$, one has:

$$(X_k^{-\varepsilon}, Y_k^{-\varepsilon}) \prec (X, Y_j^{-\varepsilon}) \sim (X_k^{-\varepsilon}, Y) \prec ((X_k^{-\varepsilon/2})_h^{-\varepsilon/2}, Y) \prec (X, Y).$$

This axiom means that if an attribute in the support of both objects, i.e. possessed to a certain extent by both objects, is modified in a way that, for the object in which the attribute is "less present", the degree of belonging is decreased, then the modified objects are less similar one to another than the initial objects were. This corresponds to a strong semantic choice: it implies that the common strong presence of an attribute is preferred to a common light presence. Moreover, the axiom states that modifying both objects degrades the similarity more than changing only one of them. In particular we notice that, for the fuzzy measure $m : \mathcal{X}^2 \to \mathbb{R}$ with $m(X) = (\sum_i x_i)$, we have:
$m(Y \cap X) > m(X_k^{-\varepsilon} \cap Y_k^{-\varepsilon}) = m(X_k^{-\varepsilon} \cap Y) = m(X \cap Y_j^{-\varepsilon}) = m((X_k^{-\varepsilon/2})_h^{-\varepsilon/2} \cap Y) = m(X \cap Y) - \varepsilon$;
and
$m(X_k^{-\varepsilon} \setminus Y_k^{-\varepsilon}) + m(Y_k^{-\varepsilon} \setminus X_k^{-\varepsilon}) = m(X \setminus Y) + m(Y \setminus X) + 2\varepsilon > m(X_k^{-\varepsilon} \setminus Y) + m(Y \setminus X_k^{-\varepsilon}) = m(X \setminus Y_j^{-\varepsilon}) + m(Y_j^{-\varepsilon} \setminus X) = m(X \setminus Y) + m(Y \setminus X) + \varepsilon > m((X_k^{-\varepsilon/2})_h^{-\varepsilon/2} \setminus Y) + m(Y \setminus (X_k^{-\varepsilon/2})_h^{-\varepsilon/2}) = m(X \setminus Y) + m(Y \setminus X)$

Axiom $S5_p$ **[monotonicity]** $\forall X, Y \in \mathcal{X},\ X \neq Y$
$\forall\, k, j, r \in s_X \cap s_Y$, such that $y_k = x_j \geq 1/2\ \ x_r < y_k$ and $\forall\, \varepsilon, \eta, \gamma$ with $\varepsilon, \eta \leq y_k$, and $\varepsilon y_k = \eta y_k + \gamma x_r$ one has:
$$(X_k^{-\eta}, Y_r^{-\gamma}) \prec (X, Y_j^{-\varepsilon}) \sim (X_k^{-\varepsilon}, Y) \sim (X_k^{-\varepsilon/2}, Y_j^{-\varepsilon/2}) \prec (Y, X)$$

Considerations similar to those made for $S5_{min}$ hold, in particular it is easy to see that in this case: $m(Y \cap X) > m(X_k^{-\eta} \cap Y_r^{-\gamma}) = m(X_k^{-\varepsilon} \cap Y) = m(X \cap Y_j^{-\varepsilon}) = m(X_k^{-\varepsilon/2} \cap Y_j^{-\varepsilon/2}) = m(X \cap Y) - \varepsilon y_k$;
and
$m(X_k^{-\eta} \setminus Y_r^{-\gamma}) + m(Y_r^{-\gamma} \setminus X_k^{-\eta}) = m(Y \setminus X) + m(X \setminus Y) + 2\varepsilon y_k - \eta - \gamma > m(X_k^{-\varepsilon} \setminus Y) + m(Y \setminus X_k^{-\varepsilon}) = m(X \setminus Y_j^{-\varepsilon}) + m(Y_j^{-\varepsilon} \setminus X) = m(X_k^{-\varepsilon/2} \setminus Y_j^{-\varepsilon/2}) + m(Y_j^{-\varepsilon/2} \setminus X_k^{-\varepsilon/2}) = m(Y \setminus X) + m(Y \setminus X) + 2\varepsilon y_k - \varepsilon > m(X \setminus Y) + m(Y \setminus X)$.

Axiom $S5_L$ **[monotonicity]** $\forall X, Y \in \mathcal{X},\ X \neq Y$
$\forall\, k, h, j \in s_X \cap s_Y$, such that $x_k < y_k,\ x_r > y_r\,,\, (r = h, j)\,,\, y_s + x_s > 1$, $(s = k, j)\,,\, y_h + x_h < 1$ and $\forall\, \varepsilon$, with $\varepsilon < \min\{x_k + y_k - 1, x_h - y_h\}$, one has:

$$(X, Y_j^{\varepsilon}) \sim (X_k^{-\varepsilon}, Y) \prec ((X_k^{-\varepsilon})_h^{-\varepsilon}, Y) \prec (X, Y)$$

Considerations similar to those regarding $S5_{min}$ hold. In particular we notice that:

$m(Y \cap X) > m(X \cap Y_j^{\varepsilon}) = m(X_k^{-\varepsilon} \cap Y) = m((X_k^{-\varepsilon})_h^{-\varepsilon}, Y) = m(X \cap Y) - \varepsilon$

and

$m(X \setminus Y_j^{-\varepsilon}) + m(Y_j^{-\varepsilon} \setminus X) = m(X_k^{-\varepsilon} \setminus Y) + m(Y \setminus X_k^{-\varepsilon}) = m(Y \setminus X) + m(X \setminus Y) + \varepsilon > m((X_k^{-\varepsilon})_h^{-\varepsilon} \setminus Y) + m(Y \setminus (X_k^{-\varepsilon})_h^{-\varepsilon}) = m(X \setminus Y) + m(Y \setminus X)$.

We notice that if axioms S1–S4 hold and m is any fuzzy measure equal to the sum of membership degrees, by using the above computations and taking into account Proposition 1, it is easy to prove that any comparative similarity $\preceq$ agreeing with a similarity measure $S_{f,g,\rho}$ defined as

$$S_{f,g,\rho}(X, Y) = \frac{f(\mathbf{x})}{f(\mathbf{x}) + \rho g(\mathbf{y})} \tag{2}$$

with $\rho > 0$ and f and g non negative increasing functions (or any strictly increasing transformation of this measure) satisfies Axioms $S5_{\top}$. Thus in particular, it is satisfied by the measures belonging to the class S_ρ defined in Equation (1), in which f and g coincide with the identity function.

On the contrary similarity measures such as Ochiai measure or Kulczynski measure, or more precisely their generalization in the fuzzy environment, do not satisfy the monotonicity axioms. In [3] a weaker form of monotonicity axiom has been introduced for the crisp case. It is one of the conditions useful to characterise comparative similarities representable by a large class of similarity measures containing as particular case the Ochiai measure. It is possible to give a similar generalization of monotonicity axioms also in a fuzzy environment. We focus here on the comparative similarities agreeing with similarity measures of the class defined in Equation (1).

3 Independence Condition

In [14], a strong axiom of independence has been introduced but it is not fulfilled by the comparative (fuzzy) similarities induced by most of the similarity measures present in the literature, in particular by the class considered in this paper (as proved in [3] for the crisp case). We now introduce a weaker form of independence in which we only require that the common characteristics are independent of the totality of the characteristics present in only one element of the pair. Let us consider a fuzzy measure m, only depending on the membership values and express, as usual, $\mathbf{x_i}, \mathbf{y_i}, \mathbf{z_i}, \mathbf{w_i}$ in terms of m.

Axiom WI [weak independence]
For any 4-tuple (X_1, Y_1), (X_2, Y_2), (Z_1, W_1), (Z_2, W_2), if one of the following conditions holds
(i) $\mathbf{x_i} = \mathbf{z_i}$ for $(i = 1, 2)$, and $\mathbf{y_1} = \mathbf{y_2}$, $\mathbf{w_1} = \mathbf{w_2}$
(ii) $\mathbf{y_i} = \mathbf{w_i}$ for $(i = 1, 2)$, and $\mathbf{x_1} = \mathbf{x_2}$, $\mathbf{z_1} = \mathbf{z_2}$
then $(X_1, Y_1) \preceq (X_2, Y_2) \Leftrightarrow (Z_1, W_1) \preceq (Z_2, W_2)$.

It must be underlined that the comparative similarities representable by a similarity measure S defined by Equation (2) satisfy this axiom, and thus in particular the elements of the class $S_{f,g,\rho}$ defined by Equation (2). We prove this assertion for hypothesis (i): by trivial computation it holds that on one hand $(X_1, Y_1) \preceq (X_2, Y_2)$ iff $f(\mathbf{x_1}) \leq f(\mathbf{x_2})$, and on the other hand $(Z_1, W_1) \preceq (Z_2, W_2)$ iff $f(\mathbf{x_1}) \leq f(\mathbf{x_2})$, leading to the desired equivalence. The proof is similar for condition (ii).

The above WI axiom is formulated independently of the t-norm of reference. Nevertheless the choice of the t-norm determines the pairs which are ruled by WI. In fact it is possible to show by simple examples that, for a specific choice of t-norm, the hypotheses of condition WI are satisfied and they are not for a different choice.

4 Representation Theorem

Now we are able to prove a theorem characterising comparative similarities representable by a class of numerical measures, in the case we adopt different t-norms.

Theorem 1. *Let us consider $\mathcal{X} = [0,1]^p$, the set of all possible fuzzy subsets of a set of p characteristics, with a t-norm $\top \in \{\min, p, L\}$. Let $\preceq$ be a binary relation on $\mathcal{X}^2 \setminus \{(\underline{0}, \underline{0})\}$. The following conditions are equivalent:*

(i) $\preceq$ is a comparative fuzzy similarity satisfying axioms $S4_\top$ and $S5_\top$ and fulfilling the weak independence property WI

(ii) for the fuzzy measure $m : \mathcal{X} \to \mathbb{R}^+$ equal to $\sum_i x_i$, there exist two non negative increasing functions f and g, with $f(0) = g(0) = 0$ such that the function $S : \mathcal{X}^2 \to [0,1]$ defined by Equation (2) represents $\preceq$.

To prove the theorem, taking into account Proposition 1, we first transform monotonicity and independence axioms in terms of a fuzzy measure of intersection and difference among fuzzy subsets. Then we can prove the theorem, by using essentially the same proof as given in [3] for the crisp case.

5 Conclusion

Considering the framework of measurement theory, we established a relation between comparative similarities, i.e. binary weak order, and numerical similarity measures so that the latter represent the former, in the case of fuzzy data. We highlighted the equivalence between specific properties satisfied by the comparative similarity and a specific form of numerical similarity, containing as particular cases Jaccard, Dice, Sorensen, Anderberg and Sokal-Sneath measures. This characterisation aims at helping users of similarity to make an appropriate choice of a similarity measure when facing a particular problem: the selection should rely on the theoretical desired properties, in terms of monotonicity and independence. The desired behaviors can be expressed in terms of induced rankings, which is more compatible with the subjective view of human beings on similarity than desired behaviors imposed on its possible numerical values.

Future works aim at studying the case of other similarity measure forms, and possibly weakening the independence axiom. Another generalization perspective concerns the study of the results obtained when one replaces the sum of the fuzzy measures of $X \setminus Y$ and $Y \setminus X$ with the fuzzy measure of $(X \setminus Y) \cup (Y \setminus X)$.

References

1. Batagelj, V., Bren, M.: Comparing resemblance measures. Journal of Classification 12, 73–90 (1995)
2. Baulieu, F.B.: A classification of presence/absence based dissimilarity coefficients. Journal of Classification 6, 233–246 (1989)
3. Bouchon-Meunier, B., Coletti, G., Lesot, M.J., Rifqi, M.: Towards a Conscious Choice of a Similarity Measure: A Qualitative Point of View. In: Sossai, C., Chemello, G. (eds.) ECSQARU 2009. LNCS (LNAI), vol. 5590, pp. 542–553. Springer, Heidelberg (2009)
4. Bouchon-Meunier, B., Rifqi, M., Bothorel, S.: Towards general measures of comparison of objects. Fuzzy Sets and Systems 84, 143–153 (1996)
5. Bouchon-Meunier, B., Rifqi, M., Lesot, M.J.: Similarities in fuzzy data mining: from a cognitive view to real-world applications. In: Zurada, J., Yen, G., Wang, J. (eds.) Computational Intelligence: Research Frontiers. LNCS, vol. 5050, pp. 349–367. Springer, Heidelberg (2008)
6. Krantz, D., Luce, R., Suppes, P., Tversky, A.: Foundations of measurement, vol. I. Academic Press, London (1971)
7. Lerman, I.C.: Indice de similarité et préordonnance associée. In: Séminaire sur les ordres totaux finis, Aix-en-Provence, pp. 233–243 (1967)
8. Lesot, M.J., Rifqi, M., Benhadda, H.: Similarity measures for binary and numerical data: a survey. Intern. J. of Knowledge Engineering and Soft Data Paradigms (KESDP) 1, 63–84 (2009)
9. Omhover, J.F., Bouchon-Meunier, B.: Equivalence entre mesures de similarités floues: application à la recherche d'images par le contenu. In: 6eme Congrès Européen de Science des Systèmes (2005)
10. Omhover, J.F., Detyniecki, M., Rifqi, M., Bouchon-Meunier, B.: Image retrieval using fuzzy similarity: Measure of equivalence based on invariance in ranking. In: IEEE Int. Conf. on Fuzzy Systems (2004)
11. Rifqi, M.: Mesures de comparaison, typicalité, et classification d'objets flous: théorie et pratique. PhD thesis, University Paris 6 (1996)
12. Rissland, E.: Ai and similarity. IEEE Intelligent Systems 21, 33–49 (2006)
13. Suppes, P., Krantz, D., Luce, R., Tversky, A.: Foundations of measurement, vol. II. Academic Press, New York (1989)
14. Tversky, A.: Features of similarity. Psychological Review 84, 327–352 (1977)