Properties Analysis of Inconsistency-based Possibilistic Similarity Measures

Ilyes Jenhani
LARODEC,
Institut Supérieur
de Gestion,
Tunis, Tunisia.
ilyes.j@lycos.com

Salem Benferhat
CRIL,
Université d'Artois,
Lens, France.
benferhat@cril.univ-artois.fr

Zied Elouedi
LARODEC,
Institut Supérieur
de Gestion,
Tunis, Tunisia.
zied.elouedi@gmx.fr

Abstract

This paper deals with the problem of measuring the similarity degree between two normalized possibility distributions encoding preferences or uncertain knowledge. Many existing definitions of possibilistic similarity indexes aggregate pairwise distances between each situation in possibility distributions. This paper goes one step further, and discusses definitions of possibilistic similarity measures that include inconsistency degrees between possibility distributions. In particular, we propose a postulate-based analysis of similarity indexes which extends the basic ones that have been recently proposed in a literature.

Keywords: Possibility theory; Similarity; Inconsistency.

1 Introduction

Uncertainty and imprecision are often inherent in modeling knowledge for most real-world problems (e.g. military applications, medical diagnosis, risk analysis, group consensus opinion, etc.). Uncertainty about values of given variables (e.g. the *type* of a detected aerial object, the *temperature* of a patient, the *property_value* of a client asking for a loan, etc.) can result from some errors and hence from non-reliability (in the case of experimental measures) or from different background knowledge (in the cognitive case of agents: doctors, etc.). As a consequence, it is possible to obtain different uncertain pieces of information about a given value from different sources. Obviously, comparing these pieces of information could be of a great interest in decision making, in case-based reasoning, in performing clustering from data having some imprecise attribute values, etc.

Comparing pieces of uncertain information given by several sources has attracted a lot of attention for a long time. For instance, we can mention the well-known Euclidean and KL-divergence [13] for comparing probability distributions. Another distance has been proposed by Chan et al. [2] for bounding probabilistic belief change. Moving to belief function theory [15], several distance measures between bodies of evidence deserve to be mentioned. Some distances have been proposed as measures of performance (MOP) of identification algorithms [5] [10]. Another distance was used for the optimization of the parameters of a belief k-nearest neighbor classifier [21]. In [16], the authors proposed a distance for the quantification of errors resulting from basic probability assignment approximations.

Many contributions on measures of similarity between two given fuzzy sets have already been made [1] [3] [6] [18]. For instance, in the work by Bouchon-Meunier et al. [1], the authors proposed a similarity measure between fuzzy sets as an extension of Tversky's model on crisp sets [17]. The measure was then used to develop an image search engine. In [18], the authors have made a comparison between existing classical similarity measures for

L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (eds): Proceedings of IPMU'08, pp. 173–180
Torremolinos (Málaga), June 22–27, 2008

fuzzy sets and proposed the sameness degree which is based on fuzzy subsethood and implication operators. Moreover, in [3] and [6], the authors have proposed many fuzzy distance measures which are fuzzy versions of classical cardinality-based distances.

This paper deals with the problem of defining similarity measures between normalized possibility distributions. In [8], a basic set of properties, that any possibilistic similarity measure should satisfy, has been proposed. This set of natural properties is too minimal and is satisfied by most existing indexes. Moreover, they do not take into account the inconsistency degree between possibility distributions. In this paper, we will mainly focus on revising and extending these properties to highlight the introduction of inconsistency in measuring possibilistic similarity.

In fact, inconsistency should be considered when measuring similarity as shown by this example: Suppose that a conference chair has to select the best paper among three selected best papers (p_1, p_2, p_3) to give an award to its authors. The conference chair decides to make a second reviewing and asks two referees r_1 and r_2 to give their preferences about the papers which, in fact, will be represented in the form of possibility distributions. Let us consider these two situations:

Situation 1: The referee r_1 expresses his full satisfaction for p_3 and fully rejects p_1 and p_2 (i.e. $\pi_1(p_1) = 0$, $\pi_1(p_2) = 0$, $\pi_1(p_3) = 1$) whereas r_2 expresses his full satisfaction for p_2 and fully rejects p_1 and p_3 (i.e. $\pi_2(p_1) = 0$, $\pi_2(p_2) = 1$, $\pi_2(p_3) = 0$). Clearly, p_1 will be rejected but the chair cannot make a decision that fully fits referees' preferences.

Situation 2: The referee r_1 expresses his full satisfaction for p_1 and p_3 and fully rejects p_2 (i.e. $\pi'_1(p_1) = 1$, $\pi'_1(p_2) = 0$, $\pi'_1(p_3) = 1$) whereas r_2 expresses his full satisfaction for p_1 and p_2 and fully rejects p_3 (i.e. $\pi'_2(p_1) = 1$, $\pi'_2(p_2) = 1$, $\pi'_2(p_3) = 0$). In this case, the chair can make a decision that satisfies both reviewers since they agree that p_1 is a good paper.

The above example shows that, in some situations, distance alone is not sufficient to make a decision since the expressed preferences in both situations have the same distance. In fact, if we consider the well-known Manhattan distance (M(x,y)=$\frac{1}{n}\sum_{i=1}^{n}|x_i - y_i|$), we obtain M($\pi_1, \pi_2$)=M($\pi'_1, \pi'_2$)=2/3. Hence, we should consider an additional concept, namely, the inconsistency degree which will play a crucial role in measuring similarity between any given two possibility distributions.

The rest of the paper is organized as follows. Section 2 gives necessary background on possibility theory. Section 3 presents the six proposed basic properties that a similarity measure should satisfy. Section 4 proposes new additional properties that take into account the inconsistency degrees. Section 5 gives some derived propositions from the proposed properties. Section 6 suggests a similarity measure that generalizes the one presented in [8]. Finally, Section 7 concludes the paper.

2 Possibility Theory

Possibility theory represents a non-classical uncertainty theory, first introduced by Zadeh [20] and then developed by several authors (e.g., Dubois and Prade [4]). In this section, we will give a brief recalling on possibility theory.

Possibility distribution

Given a universe of discourse $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$, one of the fundamental concepts of possibility theory is the notion of *possibility distribution* denoted by π. π corresponds to a function which associates to each element ω_i from the universe of discourse Ω a value from a bounded and linearly ordered valuation set (L,$<$). This value is called a *possibility degree*: it encodes our knowledge on the real world. Note that, in possibility theory, the scale can be numerical (e.g. L=[0,1]): in this case we have numerical possibility degrees from the interval [0,1] and hence we are dealing with the quantitative setting of the theory. In the qualitative setting, it is the ordering between the different possible values that is important.

By convention, $\pi(\omega_i) = 1$ means that it is fully possible that ω_i is the real world, $\pi(\omega_i) = 0$ means that ω_i cannot be the real world (is impossible). Flexibility is modeled by allowing to give a possibility degree from]0,1[. In possibility theory, extreme cases of knowledge are given by:

- *Complete knowledge*: $\exists \omega_i,\ \pi(\omega_i) = 1$ *and* $\forall\ \omega_j \neq \omega_i,\ \pi(\omega_j) = 0$.

- *Total ignorance*: $\forall\ \omega_i \in \Omega,\ \pi(\omega_i) = 1$ (all values in Ω are possible).

Possibility and Necessity measures
From a possibility distribution, two dual measures can be derived: *Possibility* and *Necessity* measures. Given a possibility distribution π on the universe of discourse Ω, the corresponding possibility and necessity measures of any event $A \subseteq 2^\Omega$ are, respectively, determined by the formulas: $\Pi(A) = \max_{\omega \in A}\ \pi(\omega)$ and $N(A) = \min_{\omega \notin A}\ (1 - \pi(\omega)) = 1 - \Pi(\overline{A})$. $\Pi(A)$ evaluates at which level A is *consistent* with our knowledge represented by π while $N(A)$ evaluates at which level A is *certainly* implied by our knowledge represented by π.

Normalization
A possibility distribution π is said to be *normalized* if there exists at least one state $\omega_i \in \Omega$ which is totally possible (i.e. $\max_{\omega\in\Omega}\{\pi(\omega)\} = \pi(\omega_i)$=1). Otherwise, π is considered as sub-normalized and in this case

$$Inc(\pi) = 1 - \max_{\omega\in\Omega}\{\pi(\omega)\} \qquad (1)$$

is called the *inconsistency degree* of π. It is clear that, for normalized π, $\max_{\omega\in\Omega}\{\pi(\omega)\} = 1$, hence Inc($\pi$)=0. The measure Inc is very useful in assessing the degree of conflict between two distributions π_1 and π_2 which is given by $Inc(\pi_1 \wedge \pi_2)$. For sake of simplicity, we take the *minimum* and *product* conjunctive ($\wedge$) operators. Obviously, when $\pi_1 \wedge \pi_2$ gives a sub-normalized possibility distribution, it indicates that there is a conflict between π_1 and π_2 ($Inc(\pi_1 \wedge \pi_2) \in]0,1]$). On the other hand, when $\pi_1 \wedge \pi_2$ is normalized, there is no conflict and hence $Inc(\pi_1 \wedge \pi_2) = 0$.

Non-specificity
Possibility theory is driven by the principle of *minimum specificity*: A possibility distribution π_1 is said to be *more specific than* π_2 if and only if for each state of affairs $\omega_i \in \Omega$, $\pi_1(\omega_i) \leq \pi_2(\omega_i)$ [19]. Clearly, the more specific π, the more informative it is.

Given a permutation of the degrees of a possibility distribution $\pi = \langle \pi_{(1)},\ \pi_{(2)},\ ...,\ \pi_{(n)} \rangle$ such that $\pi_{(1)} \geq \pi_{(2)} \geq ... \geq \pi_{(n)}$, the non-specificity of a possibility distribution π, so-called *U-uncertainty* is given by: $U(\pi) = \sum_{i=2}^{n}(\pi_{(i)} - \pi_{(i+1)})\ log_2\ i\ +\ (1 - \pi_{(1)})\ log_2\ n$.

For the sake of simplicity, for the rest of the paper, a possibility distribution π on a finite set $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ will be denoted by $\pi[\pi(\omega_1), \pi(\omega_2), ..., \pi(\omega_n)]$.

3 Basic properties of a possibilistic similarity measure

The issue of comparing possibility distributions has been studied in several works. More recently, a set of basic properties has been proposed in [8]. In this section, we will briefly recall and slightly revise these properties. Note that in this paper, we only deal with normalized possibility distributions.

Let π_1 and π_2 be two possibility distributions on the same universe of discourse Ω. A possibilistic similarity measure, denoted by $s(\pi_1, \pi_2)$, should satisfy:

Property 1. Non-negativity
$s(\pi_1, \pi_2) \geq 0$.

Property 2. Symmetry
$s(\pi_1, \pi_2) = s(\pi_2, \pi_1)$.

Property 3. Upper bound and Non-degeneracy
$\forall\ \pi_i,\ s(\pi_i, \pi_i) = 1$.

Namely, identity implies full similarity. This property is weaker than the one presented in [8] which also requires the converse, namely, s(π_i, π_j)=1 iff $\pi_i = \pi_j$.

Property 4. Lower bound
If $\forall \omega_i \in \Omega$,
i) $\pi_1(\omega_i) \in \{0, 1\}$ and $\pi_2(\omega_i) \in \{0, 1\}$,
ii) and $\pi_2(\omega_i) = 1 - \pi_1(\omega_i)$ then, s(π_1, π_2)=0.

Namely, $s(\pi_1, \pi_2) = 0$ should be obtained only when we have to compare maximally contradictory possibility distributions.

Item $i)$ means that π_1 and π_2 should be binary and since we deal with normalized possibility distributions, items $i)$ and $ii)$ imply:
$iii)$ $\exists\, \omega_q \in \Omega$ $s.t.$ $\pi_1(\omega_q) = 1$
$iv)$ $\exists\, \omega_p \in \Omega$ $s.t.$ $\pi_1(\omega_p) = 0$

Property 5. Large inclusion (specificity)
If $\forall \omega_i \in \Omega$, $\pi_1(\omega_i) \leq \pi_2(\omega_i)$ and $\pi_2(\omega_i) \leq \pi_3(\omega_i)$, which by definition means that π_1 is more specific than π_2 which is in turn more specific than π_3, we obtain: s$(\pi_1, \pi_2)\geq$s(π_1, π_3).

Property 6. Permutation
Let π_1, π_2, π_3 and π_4 be four possibility distributions such that $s(\pi_1, \pi_2) > s(\pi_3, \pi_4)$. Suppose that $\forall j = 1..4$, *and* $\omega_p, \omega_q \in \Omega$, we have $\pi'_j(\omega_p) = \pi_j(\omega_q)$, $\pi'_j(\omega_q) = \pi_j(\omega_p)$ and $\forall \omega_r \neq \omega_p, \omega_q$, $\pi'_j(\omega_r) = \pi_j(\omega_r)$. Then, $s(\pi'_1, \pi'_2) > s(\pi'_3, \pi'_4)$.

These six properties can be viewed as basic properties of any possibilistic similarity measure. They are satisfied by the following similarity measures:
Manhattan Distance:
$S_M(\pi_1, \pi_2) = 1 - \frac{\sum_{i=1}^{n}(|\pi_1(\omega_i) - \pi_2(\omega_i)|)}{n}$
Euclidean Distance:
$S_E(\pi_1, \pi_2) = 1 - \sqrt{\frac{\sum_{i=1}^{n}(\pi_1(\omega_i) - \pi_2(\omega_i))^2}{n}}$

Clearly, the above properties do not take into account the amount of conflict between possibility distributions. In fact, if we consider again our example of the introduction, where π_1=[0 0 1], π_2=[0 1 0], π'_1=[1 0 1] and π'_2=[1 1 0], then $S_M(\pi_1, \pi_2) = S_M(\pi'_1, \pi'_2) = 0.33$, $S_E(\pi_1, \pi_2) = S_E(\pi'_1, \pi'_2) = 0.18$.

To overcome this drawback, we will enrich the proposed properties by some additional ones.

4 Additional possibilistic similarity properties

The first extension concerns Property 5, where we consider a particular case of strict similarity in case of strict inclusion:

Property 7. Strict inclusion
$\forall \pi_1$, π_2, π_3 s.t. $\pi_1 \neq \pi_2 \neq \pi_3$, if $\pi_1 \leq \pi_2 \leq \pi_3$, then s$(\pi_1, \pi_2) >$ s(π_1, π_3).

Note that $\pi_1 \neq \pi_2$ and $\pi_1 \leq \pi_2$ implies $\pi_1 < \pi_2$ (strict specificity).

Next property says that, giving two possibility distributions π_1 and π_2, enhancing the degree of a given situation (with the same value) results in an increasing of the similarity between the two distributions. The similarity will be even larger, if the enhancement leads to a decrease of the amount of conflict. More precisely:

Property 8. Degree Enhancement
Let π_1 and π_2 be two possibility distributions. Let $\omega_i \in \Omega$. Let π'_1 and π'_2 s.t.:
i) $\forall j \neq i$, $\pi'_1(\omega_j) = \pi_1(\omega_j)$ and $\pi'_2(\omega_j) = \pi_2(\omega_j)$,
ii) Let α s.t. $\alpha \leq 1 - max(\pi_1(\omega_i), \pi_2(\omega_i))$.
If $\pi'_1(\omega_i) = \pi_1(\omega_i) + \alpha$ and $\pi'_2(\omega_i) = \pi_2(\omega_i) + \alpha$, then:

- If Inc$(\pi_1 \wedge \pi_2)$=Inc$(\pi'_1 \wedge \pi'_2)$, then $s(\pi_1, \pi_2) = s(\pi'_1, \pi'_2)$.
- If Inc$(\pi'_1 \wedge \pi'_2)$<Inc$(\pi_1 \wedge \pi_2)$, then $s(\pi'_1, \pi'_2) > s(\pi_1, \pi_2)$.

The intuition behind the two below properties is the following: consider two experts who provide possibility distributions π_1 and π_2. Assume that there exists a situation ω where they disagree. Now, assume that the second expert changes its mind and sets $\pi_2(\omega)$ to be equal to $\pi_1(\omega)$. Then the new similarity between π_1 and π_2 increases. This is the aim of Property 9. Property 10, goes one step further and concerns the situation when the new degree of $\pi_2(\omega)$ becomes closer to $\pi_1(\omega)$.

Property 9. Mutual convergence
Let π_1 and π_2 be two possibility distributions s.t. for some ω_i, we have $\pi_1(\omega_i) \neq \pi_2(\omega_i)$. Let π'_2 s.t.:
i) $\pi'_2(\omega_i) = \pi_1(\omega_i)$,
ii) and $\forall j \neq i$, $\pi'_2(\omega_j) = \pi_2(\omega_j)$
Hence, we obtain: $s(\pi_1, \pi'_2) > s(\pi_1, \pi_2)$.

Property 10. Generalized mutual convergence
Let π_1 and π_2 be two possibility distributions s.t. for some ω_i, we have $\pi_1(\omega_i) > \pi_2(\omega_i)$. Let π'_2 s.t.:
i) $\pi'_2(\omega_i) \in]\pi_2(\omega_i), \pi_1(\omega_i)]$,

ii) and $\forall j \neq i,\ \pi_2'(\omega_j) = \pi_2(\omega_j)$
Hence, we obtain: $s(\pi_1, \pi_2') > s(\pi_1, \pi_2)$.

Property 11 means that if one starts with a possibility distribution π_1, and modify it by decreasing (resp. increasing) only one situation ω_i (leading to π_2), or starts with a same distribution π_1 and only modify, identically, another situation ω_k (leading to π_3), then the similarity degree between π_1 and π_2 is the same as between π_1 and π_3.

Property 11. Indifference preserving
Let π_1 be a possibility distribution and α a positive number. Let π_2 s.t. $\pi_2(\omega_i) = \pi_1(\omega_i) - \alpha$ (resp. $\pi_2(\omega_i) = \pi_1(\omega_i) + \alpha$) and $\forall j \neq i$, $\pi_2(\omega_j) = \pi_1(\omega_j)$.
Let π_3 s.t. for $k \neq i$, $\pi_3(\omega_k) = \pi_1(\omega_k) - \alpha$ (resp. $\pi_3(\omega_k) = \pi_1(\omega_k) + \alpha$) and $\forall j \neq k$, $\pi_3(\omega_j) = \pi_1(\omega_j)$,
Then: s(π_1, π_2)=s(π_1, π_3).

Property 12 says that, if we consider two possibility distributions π_1 and π_2. If we increase (resp. decrease) one situation ω_p of π_1 with a degree α (leading to π_1') and, similarly, increase (resp. decrease) one situation ω_q but this time of π_2 with the same degree α (leading to π_2'), then the similarity degree between π_1 and π_1' will be equal to the one between π_2 and π_2'.

Property 12. Maintaining similarity
Let π_1 and π_2 be two possibility distributions. Let π_1' and π_2' s.t.
i) $\forall j \neq p,\ \pi_1'(\omega_j) = \pi_1(\omega_j)$ and $\pi_1'(\omega_p) = \pi_1(\omega_p) + \alpha$ (resp. $\pi_1'(\omega_p) = \pi_1(\omega_p) - \alpha$).
ii) $\forall j \neq q,\ \pi_2'(\omega_j) = \pi_2(\omega_j)$ and $\pi_2'(\omega_q) = \pi_2(\omega_q) + \alpha$ (resp. $\pi_2'(\omega_q) = \pi_2(\omega_q) - \alpha$).
Then: s(π_1, π_1')=s(π_2, π_2').

5 Derived propositions

In what follows, we will derive some propositions from the above defined properties that should characterize any possibilistic similarity measure. A consequence of Property 7 is that only identity between two distributions imply full similarity, namely:

Proposition 1 *Let s a possibilistic similarity measure s.t. s satisfies Properties 1-12. Then, $\forall \pi_i,\ \pi_j,\ s(\pi_i, \pi_j)$=1 iff π_i=π_j.*

This also means that: $\forall \pi_j \neq \pi_i,\ s(\pi_i, \pi_i) > s(\pi_i, \pi_j)$.

Besides, only completely contradictory possibility distributions imply a similarity degree equal to 0:

Proposition 2 *Let s a possibilistic similarity measure s.t. s satisfies Properties 1-12. Then, $\forall \pi_i,\ \pi_j,\ s(\pi_i, \pi_j)$=0 iff $\forall \omega_i \in \Omega$,*
i) $\pi_1(\omega_i) \in \{0, 1\}$ and $\pi_2(\omega_i) \in \{0, 1\}$,
ii) and $\pi_2(\omega_i) = 1 - \pi_1(\omega_i)$

As a consequence of Property 8, discounting the possibility degree of a same situation leads to a decrease of similarity:

Proposition 3 *Let s a possibilistic similarity measure satisfying Properties 1-12. Let π_1 and π_2 be two possibility distributions. Let $\omega_i \in \Omega$. Let π_1' and π_2' s.t.:*
i) $\forall j \neq i,\ \pi_1'(\omega_j) = \pi_1(\omega_j)$ and $\pi_2'(\omega_j) = \pi_2(\omega_j)$,
ii) Let β s.t. $\beta \leq min(\pi_1(\omega_i), \pi_2(\omega_i))$.
If $\pi_1'(\omega_i) = \pi_1(\omega_i) - \beta$ and $\pi_2'(\omega_i) = \pi_2(\omega_i) - \beta$. Then:
If Inc$(\pi_1 \wedge \pi_2)$=Inc$(\pi_1' \wedge \pi_2')$, then $s(\pi_1, \pi_2) = s(\pi_1', \pi_2')$.
If Inc$(\pi_1 \wedge \pi_2)$<Inc$(\pi_1' \wedge \pi_2')$, then $s(\pi_1 \wedge \pi_2) > s(\pi_1' \wedge \pi_2')$.

As a consequence of Property 9 and Property 10, starting from a possibility distribution π_1, we can define a set of possibility distributions that, gradually, converge to the most similar possibility distribution to π_1:

Proposition 4 *Let s a possibilistic similarity measure satisfying Properties 1-12. Let π_1 and π_2 be two possibility distributions s.t. for some ω_i, $\pi_1(\omega_i) > \pi_2(\omega_i)$. Let π_k (k=3..n) be a set of n possibility distributions. Each π_k is derived in step k from π_{k-1} as follows:*
i) $\pi_k(\omega_i) = \pi_{k-1}(\omega_i) + \alpha$
with $\alpha \in]0, \pi_1(\omega_i) - \pi_{k-1}(\omega_i)]$
ii) and $\forall j \neq i,\ \pi_k(\omega_j) = \pi_{k-1}(\omega_j)$
Hence, we obtain $s(\pi_1, \pi_2) < s(\pi_1, \pi_3) < s(\pi_1, \pi_4) < ... < s(\pi_1, \pi_n) \leq 1$.

6 An example of a similarity measure

This section proposes to analyze an extension of the Information Affinity measure, recently proposed in [8] and denoted $InfoAff$. Let us recall that $InfoAff$ takes into account the Manhattan distance ($M(\pi_1,\pi_2) = \frac{1}{n}\sum_{i=1}^{n}|\pi_1(\omega_i)-\pi_2(\omega_i)|$), along with the well known inconsistency measure. By extension, we mean that we do not restrict ourselves to the Manhattan distance, but we can also consider the Euclidean distance ($E(\pi_1,\pi_2) = \sqrt{\frac{\sum_{i=1}^{n}(\pi_1(\omega_i)-\pi_2(\omega_i))^2}{n}}$). Moreover, for the Inconsistency measure (Equation(1)), we can also take either the minimum or the product conjunctive operators.

Definition 1 *Let π_1 and π_2 be two possibility distributions on the same universe of discourse Ω. We define a measure $GAff(\pi_1,\pi_2)$ as follows:*

$$GAff(\pi_1,\pi_2) = 1 - \frac{\kappa * d(\pi_1,\pi_2) + \lambda * Inc(\pi_1 \wedge \pi_2)}{\kappa + \lambda} \tag{2}$$

where $\kappa>0$ and $\lambda>0$. d represents a (Manhattan or Euclidean) normalized metric distance between π_1 and π_2. $Inc(\pi_1 \wedge \pi_2)$ is the inconsistency degree between the two distributions (see Equation (1)) where $\wedge$ is taken as the product or min operators.

Proposition 5 *The $GAff$ measure satisfies all the proposed properties.*

Example 1 *Let us give an example to explain the proposed properties. For this example, we will take d as the Manhattan distance, $\wedge$ as the minimum conjunctive operator and $\kappa = \lambda = 1$.*

Property 7. Strict inclusion

Let $\pi_1[0.3,0.3,1]$,$\pi_2[0.6,0.3,1]$ and $\pi_3[1,0.3,1]$. Clearly $\pi_1 \leq \pi_2 \leq \pi_3$ and $\pi_1(\omega_1) < \pi_2(\omega_1) < \pi_3(\omega_1) \Rightarrow GAff(\pi_1,\pi_2) = 0.95 > GAff(\pi_1,\pi_3) = 0.88$

Property 8. Degree enhancement

Let $\pi_4[0,0,1]$, $\pi'_4[0.6,0,1]$, $\pi_5[0,1,0]$ and $\pi'_5[0.6,1,0]$ (we added 0.6 to ω_1).
We have $d(\pi'_4,\pi'_5)=d(\pi_4,\pi_5)=0.66$. But $Inc(\pi'_4 \wedge \pi'_5)=0.4 \neq Inc(\pi_4 \wedge \pi_5)=1$.
$\Rightarrow GAff(\pi'_4,\pi'_5)=0.46>GAff(\pi_4,\pi_5)=0.17$

Property 9 and 10. Mutual convergence

Let $\pi_6[0.2,1,0.5]$ and $\pi'_6[0.2,1,1]$ (We took $\pi'_6(\omega_3)=\pi_1(\omega_3)=1$) $\Rightarrow GAff(\pi_1,\pi'_6)=0.86>GAff(\pi_1,\pi_6)=0.53$.

Property 11. Indifference preserving

Let $\pi_{11}[1\ 0.8\ 0.4]$, $\alpha = 0.4$. If we subtract 0.4 from ω_2 in π_{11} or from ω_3 in $\pi_{11} \Rightarrow \pi_{12}[1\ 0.4\ 0.4]$ and $\pi_{13}[1\ 0.8\ 0]$.
$\Rightarrow GAff(\pi_{11},\pi_{12})=GAff(\pi_{11},\pi_{13})=0.93$.
If we add 0.2 to ω_2 in π_{11} or to ω_3 in $\pi_{11} \Rightarrow \pi'_{12}[1\ 1\ 0.4]$ and $\pi'_{13}[1\ 0.8\ 0.6]$.
$\Rightarrow GAff(\pi_{11},\pi'_{12})=GAff(\pi_{11},\pi'_{13})=0.96$.

Property 12. Maintaining similarity

Let $\pi_{14}[1\ 0.7\ 0]$, $\pi_{15}[1\ 0.2\ 0.7]$. If we add $\alpha = 0.3$ to ω_2 in π_{14} and to ω_3 in $\pi_{15} \Rightarrow \pi'_{14}[1\ 1\ 0]$ and $\pi'_{15}[1\ 0.2\ 1]$.
$\Rightarrow GAff(\pi_{14},\pi'_{14})=GAff(\pi_{15},\pi'_{15})=0.95$.
If we subtract $\alpha = 0.5$ to ω_2 in π_{14} and to ω_3 in $\pi_{15} \Rightarrow \pi''_{14}[1\ 0.2\ 0]$ and $\pi''_{15}[1\ 0.2\ 0.2]$.
$\Rightarrow GAff(\pi_{14},\pi''_{14})=GAff(\pi_{15},\pi''_{15})=0.91$.

Example 2 *If we reconsider the example of the referees where $\pi_1=[0\ 0\ 1]$, $\pi_2=[0\ 1\ 0]$, $\pi'_1=[1\ 0\ 1]$ and $\pi'_2=[1\ 1\ 0]$. If we apply GAff, we obtain: $GAff(\pi_1,\pi_2)=0.16 < GAff(\pi'_1,\pi'_2)=0.66$*

7 Conclusion

This paper revised and extended recently proposed properties [8] that a similarity measure between possibility distributions should satisfy. Although the Manhattan and Euclidean distances satisfy all the six basic properties, they do not satisfy the new extended ones (as shown by the example at the end of Section 3). Moreover, we have proposed a measure, namely, the *Generalized Affinity* function which satisfies all the axioms. We argue that the proposed measure is useful in many applications where uncertainty is represented by possibility distributions e.g. similarity-based possibilistic decision trees [9]. We can also mention the possibilistic clustering problem [11] which generally uses fuzzy similarity measures.

Appendix A. Proofs

For lack of space, we only provide the proof of

Proposition 5, only when d=Manhattan distance and $\wedge$=min. We can easily check that d can be replaced by the Euclidean distance and $\wedge$ by the product. Moreover, since GAff generalizes InfoAff [8], proofs of unchanged properties (Property 1, Property 2, Property 5 and Property 6) are immediate and consequently are not provided. The proof of Proposition 5 shows that our proposed measure satisfies all the proposed properties.

Proof of Proposition 5

Let us begin by showing that GAff satisfies the strong Upper and Lower bound properties derived respectively in Proposition 1 and Proposition 2.

Proposition 1:

One direction is evident since π_i=π_j $\Rightarrow$ GAff(π_i, π_j)=1 (Property 3).

Now, suppose that GAff(π_i, π_j)=1 and $\pi_i \neq \pi_j$.

GAff(π_i, π_j)=1 $\Rightarrow$ d(π_i, π_j)=0 AND Inc($\pi_i \wedge \pi_j$)=0 (since we deal with normalized distributions) $\Rightarrow$ π_i=π_j (contradiction with the assumption). Hence, GAff(π_i, π_j)=1 iff π_i=π_j.

Proposition 2:

One direction is evident since π_1=$1-\pi_2$ (with π_1 and π_2 are binary normalized possibility distributions) $\Rightarrow$ GAff(π_1, π_2)=0 (Property 4).

Now, suppose that:

i) GAff(π_1, π_2)=0 and

ii)$\pi_1 \neq 1 - \pi_2$ and

iii)π_1 and π_2 are not binary.

GAff(π_1, π_2)=0 $\Rightarrow$ $\frac{\kappa * d(\pi_1,\pi_2)+\lambda * Inc(\pi_1 \wedge \pi_2)}{\kappa+\lambda}$=1

$\Rightarrow \kappa * d(\pi_1, \pi_2) + \lambda * Inc(\pi_1 \wedge \pi_2) = \kappa + \lambda$.

Since, κ>0, λ>0 $\Rightarrow$ $d(\pi_1, \pi_2) = 1$ AND $Inc(\pi_1 \wedge \pi_2) = 1 \Rightarrow \forall\omega_i$, $|\pi_1(\omega_i) - \pi_2(\omega_i)|$=1 AND $\forall\omega_i$, min($\pi_1(\omega_i), \pi_2(\omega_i)$)=0 $\Rightarrow$

1) $\forall$ i, $\pi_1(\omega_i) \in$ {0,1} and $\pi_2(\omega_i) \in$ {0,1} and

2) $\forall$ i, $\pi_1(\omega_i)$=$1 - \pi_2(\omega_i)$ (contradiction with ii) and iii) of the above assumption).

Proofs of Property 1, Property 2, Property 5 and Property 6 are immediate since both d and Inc satisfy them as shown in [8]. Let us now prove that GAff satisfies Property 7-Property 12.

Property 7:

If π_1 is more specific than π_2 which is in turn more specific then π_3, since $\exists$ ω_0 s.t. $\pi_1(\omega_0) < \pi_2(\omega_0) < \pi_3(\omega_0)$:

$\Rightarrow$ d(π_1, π_2)<d(π_1, π_3) (hence, $\kappa * d(\pi_1, \pi_2) < \kappa * d(\pi_1, \pi_3)$) and

$\Rightarrow$ max($\pi_1 \wedge \pi_2$)=max($\pi_1 \wedge \pi_3$)=1

$\Rightarrow Inc(\pi_1 \wedge \pi_2) = Inc(\pi_1 \wedge \pi_3) = 0$

$\Rightarrow 1 - \frac{\kappa * d(\pi_1,\pi_2)+\lambda * Inc(\pi_1 \wedge \pi_2)}{\kappa+\lambda} > 1 - \frac{\kappa * d(\pi_1,\pi_3)+\lambda * Inc(\pi_1 \wedge \pi_3)}{\kappa+\lambda}$

$\Rightarrow GAff(\pi_1, \pi_2) > GAff(\pi_1, \pi_3)$.

Property 8:

We have d(π'_1, π'_2)=d(π_1, π_2) since we added the same value α to the same ω_i in π_1 and π_2. In the other hand, if min($\pi'_1(\omega_i), \pi'_2(\omega_i)$) < ($max(\pi_1 \wedge \pi_2)$) then Inc($\pi'_1 \wedge \pi'_2$) > Inc($\pi_1 \wedge \pi_2$) $\Rightarrow GAff(\pi_1, \pi_2) > GAff(\pi'_1, \pi'_2)$.

Else Inc($\pi'_1 \wedge \pi'_2$)=Inc($\pi_1 \wedge \pi_2$)

$\Rightarrow GAff(\pi'_1, \pi'_2) = GAff(\pi_1, \pi_2)$

Property 9 & 10:

We have, $\pi_2(\omega_i) \neq \pi_1(\omega_i)$ and $\forall j \neq i$, $\pi'_2(\omega_j) = \pi_2(\omega_j)$. When taking $\pi'_2(\omega_i) = \pi_1(\omega_i)$ or $\pi'_2(\omega_i) = x$ s.t. $x \in]\pi_2(\omega_i), \pi_1(\omega_i)]$, we certainly obtain:

$\Rightarrow d(\pi_1, \pi'_2) < d(\pi_1, \pi_2)$ and $Inc(\pi_1 \wedge \pi'_2) \leq Inc(\pi_1 \wedge \pi_2)$

$\Rightarrow \kappa * d(\pi_1, \pi'_2) + \lambda * Inc(\pi_1 \wedge \pi'_2) < \kappa * d(\pi_1, \pi_2) + \lambda * Inc(\pi_1 \wedge \pi_2)$

$\Rightarrow GAff(\pi_1, \pi'_2) > GAff(\pi_1, \pi_2)$

Property 11:

1) If we add α to $\pi_1(\omega_i)$ (which leads to π_2) or α to $\pi_1(\omega_j)$ (which leads to π_3) $\Rightarrow$ d(π_1, π_2)=d(π_1, π_3)=$\frac{\alpha}{|\Omega|}$ ($|\Omega|$ is the cardinality of the universe of discourse). Besides, Inc($\pi_1 \wedge \pi_2$)=Inc($\pi_1 \wedge \pi_3$)=0 (since we only deal with normalized distributions)

$\Rightarrow$ GAff(π_1, π_2)=GAff(π_1, π_3).

2) The second proof is immediate from 1) if we subtract α.

Property 12:

1) Similarly to the above proof, if we add α to $\pi_1(\omega_i)$ and keep the other degrees unchanged (which leads to π'_1) and α to $\pi_2(\omega_j)$ and keep the other degrees unchanged (which leads to π'_2) $\Rightarrow$ d(π_1, π'_1)=d(π_2, π'_2)=$\frac{\alpha}{|\Omega|}$ and Inc($\pi_1 \wedge \pi'_1$)=Inc($\pi_2 \wedge \pi'_2$)=0 (since we only deal with normalized distributions)

$\Rightarrow$ GAff(π_1, π'_1)=GAff(π_2, π'_2).

2) The second proof is immediate from 1) if we subtract α.

References

[1] B. Bouchon-Meunier, M. Rifqi and S. Bothorel: Towards general measures of comparison of objects, *Fuzzy sets and systems*, 143-153, 84(2), 1996.

[2] H. Chan and A. Darwiche: A distance measure for bounding probabilistic belief change, *International Journal of Approximate Reasoning*, 38, 149-174, 2005.

[3] B. De Baets and H. De Meyer: Transitivity preserving fuzzification schemes for cardinality-based similarity measures, *European Journal of Operational Research*, 160, 1, 726-740, 2005.

[4] D. Dubois and H. Prade: Possibility theory: An approach to computerized processing of uncertainty, *Plenum Press*, New York, 1988.

[5] D. Fixsen and R. P. S. Mahler: The modified Dempster-Shafer approach to classification, *IEEE. Trans. Syst. Man and Cybern.*, 27, 96-104, 1997.

[6] L.A. Fono, H. Gwet and B. Bouchon-Meunier: Fuzzy implication operators for difference operations for fuzzy sets and cardinality-based measures of comparison, *European Journal of Operational Research*, 183, 314-326, 2007.

[7] M. Higashi and G. J. Klir: On the notion of distance representing information closeness: Possibility and probability distributions, *International Journal of General Systems*, 9, 103-115, 1983.

[8] I. Jenhani, N. Ben Amor, Z. Elouedi, S. Benferhat and K. Mellouli: Information Affinity: a new similarity measure for possibilistic uncertain information, *In Proceedings of ECSQARU'07* , Hammamet, Tunisia, 2007, LNAI 4724, 840-852, Springer Verlag.

[9] I. Jenhani, N. Ben Amor, S. Benferhat and Z. Elouedi: Sim-PDT: a similarity-based possibilistic decision tree approach, *In Proceedings of FoIKS'08*, Pisa, Italy, 2008, LNCS 4932, 348-364, Springer Verlag.

[10] A.L. Jousselme, D. Grenier and E. Bossé: A new distance between two bodies of evidence, *Information Fusion*, 2:91-101, 2001.

[11] R. Krishnapuram and J.M. Keller: A possibilistic approach to clustering. *IEEE Transactions on Fuzzy Systems*, 1(2):98-110, 1993.

[12] T. Kroupa: Measure of divergence of possibility measures, *In Proceedings of the 6th Workshop on Uncertainty Processing*, Prague, pp. 173-181, 2003.

[13] S. Kullback and R. A. Leibler: On information and sufficiency, *Annals of Mathematical Statistics*, 22, 79-86, 1951.

[14] R. Sanguesa, J. Cabos and U. Cortes: Possibilistic conditional independence: a similarity based measure and its application to causal network learning, *International Journal of Approximate Reasoning*, 1997.

[15] G. Shafer: A mathematical theory of evidence, Princeton University Press, 1976.

[16] B. Tessem: Approximations for efficient computation in the theory of evidence, *Artificial Intelligence*, 61, 315-329, 1993.

[17] A. Tversky: Features of similarity, *Psychological Review*, 327-352, 84, 1977.

[18] X. Wang, B. De Baets and E. Kerre: A comparative study of similarity measures, *Fuzzy Sets and Systems*, Volume 73, Issue 2, 259-268, 1995.

[19] R. R. Yager. On the specificity of a possibility distribution, *Fuzzy Sets and Systems*, 50, 1992, 279-292.

[20] L. A. Zadeh: Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets ans Systems*, 1, 3-28, 1978.

[21] L. M. Zouhal and T. Denoeux: An evidence-theoric k-NN rule with paprameter optimization, *IEEE Trans. Syst. Man Cybern.*, C 28(2), 263-271, 1998.