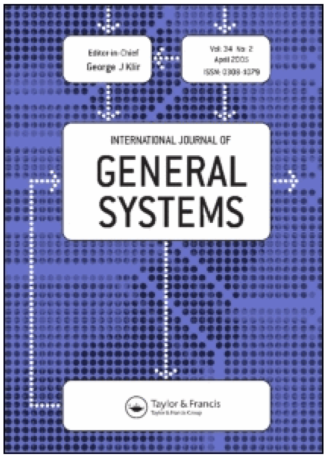

International Journal of General Systems
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713642931

COMPLETING A TOTAL UNCERTAINTY MEASURE IN THE DEMPSTER-SHAFER THEORY

Joaquín Abellán [a]; Serafín Moral [b]
[a] Departamento de Lenguajes y Sistemas, Universidad de Cadiz. 11071 Spain
[b] Departatnenio de Ciencias de la Compulacidn e Inteligencia Artificial, Universidad de Granada. Granada. Spain

Int. J. General Systems, Vol. 28(4–5), pp. 299–314
Reprints available directly from the publisher
Photocopying permitted by license only

COMPLETING A TOTAL UNCERTAINTY MEASURE IN THE DEMPSTER–SHAFER THEORY

JOAQUÍN ABELLÁN[a,*] and SERAFÍN MORAL[b]

[a]*Departamento de Lenguajes y Sistemas, Universidad de Cádiz, 11071-Cádiz, Spain;* [b]*Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada, 18071-Granada, Spain*

(Received 15 March 1997; Revised 20 September 1998; In final form 13 January 1999)

This paper seeks to complete the models that have been introduced over the last few years to quantify the uncertainty in Dempster–Shafer Theory. We examine Ichihashi and Maeda's model and we try to extend it with a correction factor. The factor discriminates between situations that having a clear difference from an intuitive point of view, the Ichihashi and Maeda's measure assigns identical values to them.

Keywords: Theory of evidence; imprecise probabilities; uncertainty; entropy; randomness; specificity

1 INTRODUCTION

It is well known that Shannon's measure of entropy plays an important role in the field of Information Theory. Over the last few years several researchers have been looking for another measure within Dempster–Shafer's Theory of Evidence (DSTE) to play the role of Shannon's entropy for probabilities. Most researchers think a basic probability assignment, b.p.a., involves two parts of uncertainty to quantify, "randomness" and "non-specificity". Furthermore, it should satisfy, Maeda and Ichihashi (1993), the following fundamental properties: it coincides with the Shannon's entropy for probabilities, it

* Corresponding author.

attains its maximum for the total ignorance and it is monotonous with respect to the inclusion of b.p.a.

Some people have proposed that it should have a range between 0 and $\ln|X|$ as in Shannon's entropy and Hartley measure, Hartley (1928), with $|X|$ being the cardinal of Universal X (Harmanec and Klir, 1994; Vejnarová and Klir, 1993). But, if we consider that the ignorance has a greater degree of uncertainty than the uniform distribution, then this range must be amplified. Other important properties have been also proposed such as *additivity* and *subadditivity*.

Dubois and Prade (1984) proposed a measure of non-specificity well accepted until now. Lamata and Moral (1987) proposed a composition of two uncertainty measures to obtain a global one. Other researchers such as Klir and Ramer (1993), Klir and Folger (1993), Vejnarová and Klir (1993), *etc.* introduce other measures of Total Uncertainty. All of them are based on Dubois and Prade's definition for measuring the non-specificity. Finally, Maeda and Ichihashi (1993) introduce a measure that seems to satisfy all the proposed properties. However, in this paper we identify a concrete situation in which it does not have a very intuitive behaviour. The goal of this paper is to introduce a factor to be added to Maeda and Ichihashi's measure to correct this problem.

This paper begins with a section on definitions in the Dempster–Shafer Theory. In Section 3 we give an example in which Maeda and Ichihashi's function does not discriminate between two situations which are clearly different from an informational point of view. The Section 4 introduces a function to correct Maeda and Ichihashi's measure. Its properties are studied. Finally, Section 5 is devoted to the conclusions.

2 BASIC DEFINITIONS

Let X be a finite set considered as a set of possible situations, $|X| = n$, $\wp(X)$ the power set of X, and x any element in X.

The Dempster–Shafer Theory, Dempster (1967) and Shafer (1976), is based on an application:

$$m: \wp(X) \rightarrow [0, 1]$$

such that $m(\emptyset) = 0$ and $\sum_{A \subseteq X} m(A) = 1$.

This function is called a *basic probability assignment* (b.p.a.). The value $m(A)$ represents the degree of belief that a specific element of X belongs to set A, but not to any particular subset of A.

The elements A of X such that $m(A) \neq 0$ are called focal elements.

There are two functions associated with each b.p.a., a belief function, Bel, and a plausibility function, Pl:

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(B)$$

$$\mathrm{Pl}(A) = \sum_{A \cap B \neq \emptyset} m(B)$$

We note that belief and plausibility are interrelated for all $A \in \wp(X)$, by

$$\mathrm{Pl}(A) = 1 - \mathrm{Bel}(\bar{A})$$

where $\bar{A}$ denotes the complement of A. Furthermore,

$$\mathrm{Bel}(A) \leq \mathrm{Pl}(A).$$

Measurement of uncertainty was first conceived in terms of classical set theory. When a b.p.a.m. focuses on a single set, *i.e.* $m(A) = 1$ and $m(B) = 0$ if $B \neq A$, then the uncertainty contained in m must collapse to the Hartley measure of set theory uncertainty, Hartley (1928), and its value is $\ln |A|$.

The classical measure of entropy, Shannon (1948), is defined by the following continuous function:

$$H(p) = -\sum_{i=1}^{n} p_i \log_a(p_i),$$

where $p\ (p_1, \ldots, p_n)$ is a probability distribution.

The non-specificity function, introduced by Dubois and Prade (1984), represents a measure of imprecision associated with a b.p.a. and has the following expression:

$$I(m) = \sum_{A \subseteq X} m(A) \ln |A|.$$

$I(m)$ attains its minimum, zero, when m is a probability distribution. The maximum, $\ln|X|$, is obtained for a b.p.a., m, with $m(X)=1$ and $m(A)=0, \forall A \subset X$.

Delgado and Moral (1987) defined a relationship order within the set of b.p.a. on a finite universal, X:

DEFINITION 1 *Let m and m' be two b.p.a. on X, m is said to be included in m' ($m \subseteq m'$) if and only if there exists an application $t_A : \wp(A) \to [0, 1]$, for every $A \subseteq X$, satisfying*:

$$m(A) = \sum_{B|B\subseteq A} t_A(B), \quad \forall A \subseteq X,$$

$$m'(B) = \sum_{A|A\supseteq B} t_A(B), \quad \forall B \subseteq X.$$

We take the following definitions from Lamata and Moral (1987):

DEFINITION 2 *Let m be a b.p.a. on the Cartesian product $X \times Y$. The projection of m on X is defined as the b.p.a. m_X given by*

$$m_X(A) = \sum_{P_X(B)=A} m(B),$$

where $P_X(B) = \{a \in X \mid \exists b \in Y, (a,b) \in B\}$.

Similarly, we may define the projection m_Y on Y.

DEFINITION 3 *Let m be a b.p.a. on $X \times Y$ with projections m_X and m_Y. Then there is strong independence under m iff*

$$m(A \times C) = m_X(A) \cdot m_Y(C); \qquad \forall A \subset X, \quad \forall C \subset Y.$$

If m is a b.p.a. on $X \times Y$ and there is strong independence under m, then the following properties can easily be checked:

$$\mathrm{Pl}(A \times C) = \mathrm{Pl}_X(A) \cdot \mathrm{Pl}_Y(C),$$
$$\mathrm{Bel}(A \times C) = \mathrm{Bel}_X(A) \cdot \mathrm{Bel}_Y(C),$$
$$m(B) > 0 \;\Rightarrow\; \exists A \subset X, C \subset Y \quad \text{such that} \quad B = A \times C.$$

Let R be a total uncertainty measure, $R : \mathbf{B} \to [0, \infty)$, with $\mathbf{B}$ being the set of all belief measures on $X \times Y$.

DEFINITION 4 *We said R is subadditive iff, for every b.p.a. m,*

$$R(m) \leq R_X(m_X) + R_Y(m_Y),$$

where R_X and R_Y are uncertainty functions induced from R on X and Y respectively.

DEFINITION 5 *R is additive iff, for strong independence under m, it verifies that:*

$$R(m) = R_X(m_X) + R_Y(m_Y).$$

DEFINITION 6 *Let A be a subset of the vectorial space $\mathbf{R}^n$. We denote by* Fr(*A*) *the frontier set of A:*

$$\mathrm{Fr}(A) = \{a \in \mathbf{R}^n \mid B(a,\partial) \cap A \neq \emptyset \wedge B(a,\partial) \cap \bar{A} \neq \emptyset, \forall \partial > 0, \partial \in \mathbf{R}\},$$

where $B(a, \partial) = \{b \in \mathbf{R}^n \mid d(a,b) \leq \partial\}$ and $d(a,b)$ is a distance function on $\mathbf{R}^n$.

3 MAEDA AND ICHIHASHI'S FUNCTION

Maeda and Ichihashi (1993) proposed an uncertainty function. It quantified the randomness and non-specificity contained in a b.p.a. on X. The function is:

$$\mathrm{UT}(m) = I(m) + G(m),$$

where $I(m)$ is Dubois and Prade's non-specificity function and $G(m)$ is the solution of the problem:

$$\mathrm{Max}\left\{ -\sum_{x \in X} p_x \ln p_x \right\},$$

where the maximum is taken over all the probability distributions on C_m, and C_m a closed convex set on $\mathbf{R}^{|X|}$, Harmanec and Klir (1994), that is defined as the set of probability distributions $\{(p_x) \mid x \in X\}$

satisfying the constraints:

(a) $p_x \in [0, 1]$ for all $x \in X$ and $\sum_{x \in X} p_x = 1$;
(b) $\text{Bel}(A) \leq \sum_{x \in A} p_x \leq 1 - \text{Bel}(\bar{A})$for all $A \subseteq X$.

It is easy to prove the UT(m) satisfies the following requirements:

(a) It coincides with the Shannon's entropy for probabilities.
(b) It reaches its maximum for the total ignorance.

And we can see in Dubois and Prade (1987) or Lamata and Moral (1987):

(c) It is monotonous with respect to the inclusion of b.p.a. ($m_2 \subseteq m_1 \Rightarrow \text{UT}(m_1) \leq \text{UT}(m_2)$.
(d) It satisfies the additivity property.
(e) It satisfies the subadditivity property.

We can also see that $G(m)$ satisfies the previous requirements and it could be considered like a total uncertainty measure, Harmanec and Klir (1994).

But, we think UT misses some aspect of uncertainty. In the following example, we see what the problem is.

Example 1 Let the following b.p.a. be on the universal $X = \{x_1, x_2, x_3\}$:

$$m \begin{cases} m_{123} = 0.4 \\ m_1 = 0.2 \\ m_2 = 0.2 \\ m_3 = 0.2 \end{cases}$$

and

$$m' \begin{cases} m'_{123} = 0.2 \\ m'_{23} = (m_{123} - m'_{123}) \ln(3)/\ln(2) \simeq 0.317 \\ m'_1 = m'_2 = m'_3 = (1 - m'_{123} - m'_{23})/3 \simeq 0.161, \end{cases}$$

where $m_i = m(\{x_i\})$, $m_{ij} = m(\{x_i, x_j\})$, $i,j \in \{1, 2, 3\}$ and $m_{123} = m(\{x_1, x_2, x_3\})$. Similarly for m'.

If we observe these functions, it is reasonable to think that m should represent more uncertainty than m', as m is completely symmetrical and m' points to $\{x_2, x_3\}$:

$$\text{Pl}(\{x_1, x_2\}) = \text{Pl}(\{x_1, x_3\}) = \text{Pl}(\{x_2, x_3\}) = 0.8,$$
$$\text{Pl}(\{x_1\}) = \text{Pl}(\{x_2\}) = \text{Pl}(\{x_3\}) = 0.6,$$
$$\text{Bel}(\{x_1, x_2\}) = \text{Bel}(\{x_1, x_3\}) = \text{Bel}(\{x_2, x_3\}) = 0.4,$$
$$\text{Bel}(\{x_1\}) = \text{Bel}(\{x_2\}) = \text{Bel}(\{x_3\}) = 0.2,$$

and

$$\text{Pl}'(\{x_1, x_2\}) = \text{Pl}'(\{x_1, x_3\}) = \text{Pl}'(\{x_2, x_3\}) = 0.839,$$
$$0.361 = \text{Pl}'(\{x_1\}) \ll \text{Pl}'(\{x_2\}) = \text{Pl}'(\{x_3\}) = 0.839,$$
$$0.322 = \text{Bel}'(\{x_1, x_2\}) = \text{Bel}'(\{x_1, x_3\}) \ll \text{Bel}'(\{x_2, x_3\}) = 0.639,$$
$$\text{Bel}'(\{x_1\}) = \text{Bel}'(\{x_2\}) = \text{Bel}'(\{x_3\}) = 0.161.$$

Each probability distribution can be represented on an equilateral triangle, De Campos (1986) and Dempster (1967), in which $p(x_i)$ is the distance to the x_i edge. The two convex sets associated to m and m' are given in Figs. 1 and 2 respectively.

Graphically, if we ignore the common points of m and m', for each point of m' representing a distribution of probabilities, there is a point in m with more uncertainty when using the Shannon measure. But, in the other way round it is not true. Intuitively, UT(m) should be greater than UT(m').

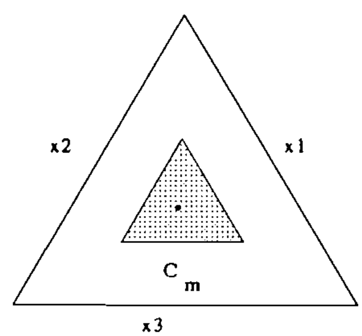

FIGURE 1

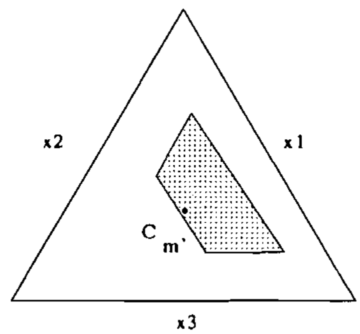

FIGURE 2

Both, m' and m, have the same non-specificity value $I(m) = I(m') = 0.439$. And $G(m) = G(m') = \ln(3)$, because $p_U \in C_m$ and $p_U \in C_{m'}$, where p_U is the uniform distribution on X. Then m and m' have the same Maeda and Ichihashi's uncertainty. In the following section we introduce the Kullback factor which will discriminate between these two situations.

4 THE KULLBACK FACTOR

Here we introduce a factor with some interesting properties, which can be used to improve Maeda and Ichihashi's measure. Our starting point will be the cross entropy between two probability distributions as introduced by Kullback (1968):

$$K(p, q) = \sum_{x \in X} p_x \ln\left(\frac{p_x}{q_x}\right),$$

where p and q are two probability distributions on a finite set X. This function is similar to an information measure and may be considered as a measure of *direct divergence*, Kullback (1968). It does not have all the properties of a distance.

We use this function in the following way. Let

$$R(m) = \underset{p \in \mathrm{Fr}(C_m)}{\mathrm{Min}} K(p, \hat{q}),$$

where q is such that $G(m) = -\sum_{x \in X} \hat{q}_x \ln(\hat{q}_x)$, *i.e.*, the probability distribution with maximum entropy inside C_m; where C_m is the convex set of probabilities associated to m (De Campos, 1986; Dempster, 1967). We call $R(m)$ the *Kullback factor of m*.

It is useful to use this function in the previous example, where $\hat{q} = p_U$ for m and m', and as we see in Figs. 1 and 2, $R(m) > R(m')$. It is easy to prove $\mathrm{Min}_{p\in \mathrm{Fr}(C_m)} K(p, \hat{q})$ is attained in the point of minimum distance to the frontier sets of C_m and $C_{m'}$ with $R(m) = 0.0437$ and $R(m') = 0.0017$.

Then we define a measure of Total Uncertainty

$$\mathrm{UTR}(m) = I(m) + G(m) + R(m).$$

Now, m has more entropy than m' as was intuitively expected.

4.1 Properties

With the preceding notation, we have the following properties.

LEMMA 1 *If* $p_U \in \mathrm{Fr}(C_m)$ *then* $R(m) = 0$.

Proof We assume that $0 \ln(0) = 0$.

Since C_m is a closed set, $p_U \in C_m$. Then $R(m) = K(p_U, p_U) = 0$.

LEMMA 2 *If* $p_U \notin C_m$ *then* $R(m) = 0$.

Proof Let $\mathrm{Max}_{p\in C_m} H(p) = H(p')$. It is only necessary to prove $p' \in \mathrm{Fr}(C_m)$.

We suppose $p' \notin \mathrm{Fr}(C_m)$. Then we choose $\alpha \in \mathbf{R}$, $\alpha \in (0, 1)$ such that

$$p'' = \alpha \cdot p_U + (1 - \alpha) \cdot p'$$

and $p'' \in C_m$.

For the continuity of H, $H(p'') > H(p')$. Hence $p' \in \mathrm{Fr}(C_m)$.

PROP 1 *$R(m)$ is well defined.*

Proof If $p_U \notin C_m$ then, using Lemma 2, $R(m) = 0$. If $p_U \in C_m$ then $R(m) = \ln(n) - H(p^*)$, for some $p^* \in \mathrm{Fr}(C_m)$.

PROP 2 *$R(m) \geq 0$, $\forall m$ b.p.a. on X finite.*

Proof Let $R(m)=\sum_{x\in X}p_x\ln(p_x/q_x)$, for a distribution of probabilities $p\in C_m$. Now, using Gibbs' inequality, Klir and Folger (1993), we have

$$-\sum_{x\in X} p_x \ln(p_x) \leq -\sum_{x\in X} p_x \ln(\hat{q}_x)$$

and $R(m)\geq 0$.

PROP 3 *If m is a probability distribution then* $R(m)=0$.

PROP 4 *R attains its maximum value for the total ignorance. Then*

$$R(m) = \ln(n) - \ln(n-1).$$

Proof With the above notation, let m be a b.p.a. representing the total ignorance on X. Then $m_X=1$.

We know that $\mathrm{Max}_{p\in \mathrm{Fr}(C_m)} H(p) = \ln(n-1)$. Then

$$R(m) = \underset{p\in \mathrm{Fr}(C_m)}{\mathrm{Min}} K(p,p_U) = \ln(n) - \underset{p\in \mathrm{Fr}(C_m)}{\mathrm{Max}} H(p) = \ln(n) - \ln(n-1).$$

Now $\forall m'$ b.p.a. on X, $R(m')\leq R(m)$.

If $p_U\notin C_{m'}$ then, using Lemma 2, $R(m')=0\leq R(m)$.

If $p_U\in C_{m'}$ we consider $p'\in \mathrm{Fr}(C_{m'})$ such that $p'=\alpha p_U+(1-\alpha)p_{U_{n-1}}$ with $\alpha\in[0,1]$, where $p_{U_{n-1}}$ is the uniform probability on some set $X'\subset X$ with $|X'|=n-1$.

Then, by the continuity of H

$$\ln(n) = H(p_U) \geq H(p') \geq H(p_{U_{n-1}}) = \ln(n-1).$$

Now, $R(m')=\ln(n)-H(p^*)$, for some $p^*\in \mathrm{Fr}(C_{m'})$ and

$$R(m') = \ln(n) - H(p^*) \leq \ln(n) - H(p') \leq \ln(n) - \ln(n-1) = R(m).$$

PROP 5 *R is a monotonous function of m.*

Proof Let m and m' be two b.p.a. such that $m'\subseteq m$ in the Definition 1 sense. Then for $p\in C_m$, distribution of probabilities, $p\in C_{m'}$.

Cases:

A. $p_U \notin C_{m'} \Rightarrow p_U \notin C_m \underset{\text{Lemma 2}}{\Rightarrow} R(m) = R(m') = 0$.
B. $p_U \in C_{m'}$:

B.1. $p_U \notin C_m \underset{\text{Lemma 2}}{\Rightarrow} R(m) = 0 \leq R(m')$.
B.2. $p_U \in C_m$. Let $p^* \in C_{m'}$ such that

$$R(m') = \sum_x p_x^* \ln\left(\frac{p_x^*}{1/n}\right) = \ln(n) - H(p^*).$$

Since $C_m \subseteq C_{m'}$ there exists $\alpha \in \mathbf{R}$, $\alpha \in [0, 1]$, such that $p' = \alpha \cdot p^* + (1 - \alpha) \cdot p_U$, and $p' \in \mathrm{Fr}(C_m)$. Then by the continuity of H, $H(p^*) \leq H(p')$ and

$$R(m) = \underset{p \in \mathrm{Fr}(C_m)}{\mathrm{Min}} [\ln(n) - H(p)] \leq \ln(n) - H(p') \leq \ln(n) - H(p^*) = R(m').$$

LEMMA 3 *Let m be a b.p.a. on $X \times Y$ with projections m_X and m_Y. Let p_{U_X} the uniform distribution for X and p_{U_Y} the uniform distribution for Y. Then*

$$p_U \in C_m \;\Rightarrow\; \begin{cases} p_{U_X} \in C_{m_X} \text{ and} \\ p_{U_Y} \in C_{m_Y}. \end{cases}$$

Proof Let $|X| = n_X$, $|Y| = n_Y$ and $n = n_X \cdot n_Y$. Then $p_{U_X} = (1/n_X, \ldots, 1/n_X)$ and $p_{U_Y} = (1/n_Y, \ldots, 1/n_Y)$.

We know if a probability distribution on X, p, such that $p(x) = \sum_{y \in Y} p_U(x, y)$ then $p \in C_{m_X}$. Now,

$$p(x) = \sum_{y \in Y} p_U(x, y) = n_Y \cdot \frac{1}{n} = \frac{n_Y}{n_X n_Y} = \frac{1}{n_X} = p_{U_X}(x), \quad \forall x \in X.$$

Idem for m_Y.

LEMMA 4 *Let m be a b.p.a. on $X \times Y$ with projections m_X and m_Y, such that there is strong independence under m. Let p_{U_X} the uniform distribution for X and p_{U_Y} the uniform distribution for Y. Then*

$$p_U \in C_m \;\Leftarrow\; \begin{cases} p_{U_X} \in C_{m_X} \text{ and} \\ p_{U_Y} \in C_{m_Y}. \end{cases}$$

Proof Using the independence hypothesis, let $p^{xy} \equiv p_{U_X} \cdot p_{U_Y} \in C_m$, (Harmanec and Klir, 1994). But

$$p^{xy}(x,y) = p_{U_X}(x) \cdot p_{U_Y}(y) = \frac{1}{n_X} \cdot \frac{1}{n_Y} = \frac{1}{n} = p_U(x,y), \quad \forall (x,y) \in X \times Y.$$

LEMMA 5 *Let m be a b.p.a. on $X \times Y$ with projections m_X and m_Y, such that there is strong independence under m. Let $p \in \mathrm{Fr}(C_{m_X})$ and $q \in \mathrm{Fr}(C_{m_Y})$. Then $pq \in \mathrm{Fr}(C_m)$.*

Proof We take the distance on $\mathbf{R}^n$: $d(u,v) = \mathrm{Max}_{i \in \{1,\ldots,n\}} |u_i - v_i|$; $u, v \in \mathbf{R}^n$. By hypothesis:

$$\begin{aligned}
\forall \partial_1 > 0 \quad & \exists p' \in \mathbf{R}^{n_X} \text{ such that } \operatorname*{Max}_x |p_x - p'_x| \le \partial_1 \text{ and } p' \in C_{m_X} \\
& \exists p'' \in \mathbf{R}^{n_X} \text{ such that } \operatorname*{Max}_x |p_x - p''_x| \le \partial_1 \text{ and } p'' \in \overline{C_{m_X}}, \\
\forall \partial_2 > 0 \quad & \exists q' \in \mathbf{R}^{n_Y} \text{ such that } \operatorname*{Max}_y |q_y - q'_y| \le \partial_2 \text{ and } q' \in C_{m_Y} \\
& \exists q'' \in \mathbf{R}^{n_Y} \text{ such that } \operatorname*{Max}_y |q_y - q''_y| \le \partial_2 \text{ and } q'' \in \overline{C_{m_Y}}.
\end{aligned}$$

Now, $\forall \partial > 0$ we take $\partial_1 = \partial_2 = \partial/2$ and since $p'q' \in C_m$ (Harmanec and Klir, 1994) and

$$pq - p'q' = q(p - p') - p'(q' - q),$$

then

$$\begin{aligned}
\operatorname*{Max}_{\substack{x\in X\\ y\in Y}} |p_x q_y - p'_x q'_y| &= \operatorname*{Max}_{\substack{x\in X\\ y\in Y}} |q_y(p_x - p'_x) - p'_x(q'_y - q_y)| \\
&\le \operatorname*{Max}_{\substack{x\in X\\ y\in Y}} [|q_y||(p_x - p'_x)| + |p'_x||(q'_y - q_y)|] \\
&\le \operatorname*{Max}_{\substack{x\in X\\ y\in Y}} [|(p_x - p'_x)| + |(q'_y - q_y)|] \\
&= \operatorname*{Max}_{x\in X} |(p_x - p'_x)| + \operatorname*{Max}_{y\in Y} |(q'_y - q_y)| \\
&\le \frac{\partial}{2} + \frac{\partial}{2} \le \partial.
\end{aligned}$$

Similarly for p'' and q'', since $p''q'' \in \overline{C_{m_X}} \times \overline{C_{m_Y}} \subseteq \overline{C_m}$.

PROP 6 *R is subadditive.*

Proof With the above notation, let m be a b.p.a. on $X \times Y$, then

$$R(m) \leq R_X(m_X) + R_Y(m_Y).$$

Cases:

A. If $p_U \notin C_m$

$$0 \underset{\text{Lemma 2}}{=} R(m) \leq R_X(m_X) + R_Y(m_Y).$$

B. If $p_U \in C_m$. Let

$$R(m) = \sum_{x,y} \hat{p}_{xy} \ln\left(\frac{\hat{p}_{xy}}{1/n}\right); \quad \hat{p} \in C_m$$

and using Lemma 3

$$R_X(m_X) = \sum_{x} \hat{p}^1_x \ln\left(\frac{\hat{p}^1_x}{1/n_X}\right); \quad \hat{p}^1 \in \text{Fr}(C_{m_X}),$$

$$R_Y(m_Y) = \sum_{y} \hat{p}^2_y \ln\left(\frac{\hat{p}^2_y}{1/n_Y}\right); \quad \hat{p}^2 \in \text{Fr}(C_{m_Y}).$$

We take $m_X \times m_Y$ b.p.a. on $X \times Y$ such that $m_X \times m_Y (A \times B) = m_X(A) m_Y(B)$, with $A \subseteq X$ and $B \subseteq Y$. Then there is strong independence under $m_X \times m_Y$ and $C_{m_X} \times C_{m_Y} \subseteq C_{m_X \times m_Y}$ (Harmanec and Klir, 1994).

By Lemma 5, $\hat{p}^1\hat{p}^2 \in \text{Fr}(C_{m_X \times m_Y})$.

Since $C_{m_X} \times C_{m_Y} \subseteq C_{m_X \times m_Y}$, a convex set, then the Convex Hull of $(C_{m_X} \times C_{m_Y})$ is also contained in $C_{m_X \times m_Y}$. Hence, $C_m \subseteq \text{CH}(C_{m_X} \times C_{m_Y}) \subseteq C_{m_X \times m_Y}$.

Let q be a distribution of probabilities such that $q \in \text{Fr}(C_m)$ and $q = \alpha p_U + (1-\alpha)\hat{p}^1\hat{p}^2$, with $\alpha \in [0, 1]$.

Now,

$$R(m) = \underset{p \in \text{Fr}(C_m)}{\text{Min}} [\ln(n) - H(p)] = \ln(n) - \underset{p \in \text{Fr}(C_m)}{\text{Max}} H(p) \leq \ln(n) - H(q)$$

By the continuity of H, $H(p_U) \geq H(q) \geq H(\hat{p}^1\hat{p}^2) = H(\hat{p}^1) + H(\hat{p}^2)$, and

$$R(m) \leq \ln(n) - H(q) \leq \ln(n_X) + \ln(n_Y) - (H(\hat{p}^1) + H(\hat{p}^2)) = R_X(m_X) + R_Y(m_Y).$$

Function R in general does not satisfy the additive property as we can see in the following example.

Example 2 We choose m_X a b.p.a. on X such that $p_{U_X} \notin m_X$ and m_Y on Y with $p_{U_Y} \in m_Y$ but $p_{U_Y} \notin \text{Fr}(C_{m_Y})$. It results that $R_X(m_X) = 0$ and $R_Y(m_Y) > 0$.

Let m_Y be equals to m in Example 1, then $R_Y(m_Y) = 0.0437$.

Let m_X be a b.p.a. on $X = \{a, b, c\}$ such that $m_X(\{a\}) = 1$ and 0 in other case. Obviously $p_{U_X}(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \notin m_X$ and $R_X(m_X) = 0$.

Now

$$\text{Bel}(\{(a, 1)\}) = 0.2 > \tfrac{1}{9}.$$

This implies that $p_U(\frac{1}{9}, \frac{1}{9}, \ldots, \frac{1}{9}) \notin C_m$ because if $p_U \in C_m$ then, by Lemma 3, $p_{U_X} \in C_{m_X}$.

Now, by Lemma 2, $R(m) = 0$ and

$$R(m) < R_X(m_X) + R_Y(m_Y).$$

The ampliation of Maeda and Ichihashi's uncertainty measure by R function, satisfies the three indispensable properties within the frame of belief functions (Maeda and Ichihashi, 1993):

- It is reduced to the Shannon entropy when a b.p.a. becomes a probability distribution,

$$R(p) = 0,$$
$$I(p) = 0,$$
$$G(p) = H(p).$$

- It is maximal for the total ignorance, represented by a b.p.a. m such that $m(X) = 1$ and $m(A) = 0, \forall A \subset X$

$$\text{UTR}(m) = 3\ln(n) - \ln(n - 1).$$

- It is monotonous with respect to random set inclusion (Prop 5).

We also proved, it is a subadditive function by Prop 6, although, generally it is not additive as we see in Example 2.

5 CONCLUSIONS

$R(m)$ is not a measure of randomness or specificity but can be a good complement for a total uncertainty measure.

The behaviour or $R(m)$ depends on whether m is in the set $S_U = \{m|p_U \in m\}$, *i.e.*, the uniform distribution is in C_m. If $m \notin S_U$, then $R(m) = 0$ and $\text{UTR}(m) = \text{UT}(m)$, but, in this case we think that Maeda and Ichihashi's measure has a correct behaviour. $R(m)$ adds a positive value to $\text{UT}(m)$ when $m \in S_U$. In this case, for the same specificity, $R(m)$ takes into account whether the uniform distribution is really in the centre of C_m or very close to the frontier set. In the former case $R(m)$ is greater than in the last. The uncertainty is greater when all distributions are around the uniform distribution. $\text{UTR}(m)$ takes this fact into account, while it is missing in $\text{UT}(m)$. However, the differences between $\text{UTR}(m)$ and $\text{UT}(m)$ are never big, being always lower than $\ln(n) - \ln(n-1)$.

If we want to quantify the uncertainty in a b.p.a., m, we think $G(m)$ could not be enough to quantify the part of randomness contained in it.

This ampliation may not be the only possible one. A function of a distance function on $\mathbf{R}^n$ could play the same role.

Finally, another issue from R is its possible application to general convex sets of probabilities. This is feasible because its definition is based on the associated convex set of probabilities and therefore its extension is immediate.

References

De Campos, L.M. (1986) Caracterización y estudio de medidas e integrales difusas a partir de probabilidades. Tesis Doctoral, Univ. Granada, Granada.

Delgado, M. and Moral, S. (1987) A definition of inclusion for evidences. *Fuzzy Mathematics*, **7**, 81–87.

Dempster, A.P. (1967) Upper and lower probabilities induced by a multi-valuated mapping. *Ann. Math. Statistic.*, **38**, 325–339.

Dubois, D. and Prade, H. (1984) A note on measure of specificity for fuzzy sets. *BUSEFAL*, **19**, 83–89.

Dubois, D. and Prade, H. (1987) Properties and measures of information in evidence and possibility theories. *Fuzzy Sets and System*, **24**(2), 183–196.

Harmanec, D. and Klir, J. (1994) Measuring total uncertainty in Dempster–Shafer theory: a novel approach. *Int. J. General System*, **22**, 405–419.

Hartley, R.V.L. (1928) Transmission of information. *The Bell Systems Technical Journal*, **7**, 535–563.

Klir, J. and Folger, T. (1993) *Fuzzy Sets, Uncertainty and Information*. Prentice-Hall, Englewood Cliffs, NJ.

Klir, J. and Ramer, A. (1993) Measures of discord in the Dempster–Shafer theory. *Information Sciences* **67**, 35–50.

Kullback, S. (1968) *Information Theory and Statistics*. Dover, New York.

Lamata, M.T. and Moral, S. (1987) Measures of entropy in the theory of evidence. *Fuzzy Sets and System*, **12**, 193–226.

Maeda, Y. and Ichihashi, H. (1993) An uncertainty measure with monotonicity under the random set inclusion. *Int. J. General System*, **21**, 379–392.

Vejnarová, J. and Klir, J. (1993) Measures of strife in Dempster–Shafer theory. *Int. J. General System*, **22**, 22–42.

Shafer, G. (1976) *A Mathematical Theory of Evidence*. Princeton University Press, Princeton.

Shannon, C.E. (1948) Entropy and specificity in a mathematical theory of evidence. *The Bell System Technical Journal*, **27**, 379–423, 623–656.

Joaquín Abellán received a B.S. degree in Mathematics from the University of Granada (Spain) in 1989. From 1994 to 1998 he was an assistant professor at the University of Cádiz (Spain). Now, he is a doctoral candidate in the Department of Computer Science and Artificial Intelligence at the University of Granada. His research interests are representation of uncertainty through the theory of evidence and imprecise probabilities.

Serafín Moral received his Ph.D. degree in 1984 from the University of Granada. He is a professor of Computer Science and Artificial Intelligence and member of the research group on Uncertainty Treatment on Artificial Intelligence (UTAI) at the University of Granada. His current research interests are imprecise probabilities, propagation algorithms in dependence graphs, and uncertain and defeasible reasoning in general.