Chaos, Solitons and Fractals 95 (2017) 195–199

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Analyzing properties of Deng entropy in the theory of evidence

Joaquín Abellán

Department of Computer Science & Artificial Intelligence, University of Granada, 18071, Granada, Spain

ARTICLE INFO

Article history:
Received 17 September 2016
Revised 4 November 2016
Accepted 20 December 2016
Available online 28 December 2016

Keywords:
Imprecise probabilities
Theory of evidence
Measures of uncertainty
Discord
Non-specificity
Deng entropy

ABSTRACT

The theory of Evidence, or Shafer-Dempster theory (DST), has been widely used in applications. The DST is based on the concept of a basic probability assignment. An important part of this theory is the quantification of the information-based-uncertainty that this function represents. A recent measure of uncertainty (or information) in this theory, called the Deng entropy, has appeared as an interesting alternative to the measures presented so far. This measure quantifies the both types of uncertainty found in DST, then it is considered as a total uncertainty measure (TU). It is shown that this measure does not verify some of the essential properties for a TU in DST such as monotonicity, additivity and subadditivity. Also, the definition of this new measure produces other debatable situations. These shortcomings call in question the utility of this measure in applications.

1. Introduction

The theory of Evidence [8,21], mostly known as Dempster–Shafer's theory (DST), was presented as an extension of the classical probability theory (PT). In the DST a new concept, called *basic probability assignment* (bpa), was introduced to generalize the one of the probability distribution in the PT.

In the 90s were presented many studies about the uncertainty based information that a bpa can represent. The majority of the measures presented had as starting point the Shannon's entropy [22]. In DST, were found more types of uncertainty than in PT, as it is logical because DST includes the PT. Yager [23] makes the distinction in DST between two types of uncertainty called: *discord* (*randomness* or *conflict*) and *non-specificity* respectively. Harmanec and Klir [11] presented a total uncertainty measure (TU) in DST, i.e. a measure that quantifies both types of uncertainty, that has been justified by an axiomatic approach (Klir and Wierman [18]). They also established, for such type of measures, a set of five desired properties that a TU must verify. Abellán and Masegosa [2] extended that set adding the important property of the monotonicity.

As far, the upper entropy measure is the only measure that verifies all the basic required properties exposed in Klir and Wierman [18] and Abellán and Masegosa [2].

Very recently, a new measure called *Deng entropy* [7] has been presented as an alternative in DST. This function considers that the degree of uncertainty is strongly related with the number of possible alternatives. But this new measure presents some shortcomings that make us to be cautious if we want to use it in applications. We will see that this measure does not verify the majority of the above mentioned properties and has other weakness motivated by its definition. It is true that some of those properties could be questionable but other ones are very important and essential for such measures. We remark that a measure of that type must take into account the decreasing or increasing in information or at least not express an erroneous situation. Also, when two non-interactive evidences can be joined, the total amount of information can be not be increased or decreased by an uncertainty measure. Finally, if an evidence on a finite set can be decomposed on more simple sets, then the total amount of information can not be increased.

The paper is organized as follows. Section 2 reviews briefly the Dempster-Shafer Theory (DST). Section 3 is devoted to show some of the most important measures of uncertainty presented in DST, and exposes the set of basic properties necessary for such measures. Section 4 studied the set of properties verified by the new measure, and shows some of the shortcoming found on that measure. Section 5 is dedicated to the conclusions and future works.

2. Dempster–Shafer theory of Evidence

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. Dempster–Shafer theory is based on the concept of basic probability assignment. A *basic probability assignment* (bpa), also called *mass assignment*, is a mapping $m: \wp(X) \to [0, 1]$, such that $m(\emptyset) = 0$ and $\sum_{A \subseteq X} m(A) = 1$. A set A such that $m(A) > 0$ is called a *focal element* of m.

E-mail address: jabellan@decsai.ugr.es

http://dx.doi.org/10.1016/j.chaos.2016.12.024

Let X, Y be finite sets. Considering the product space of possible situation $X \times Y$ and m a bpa on $X \times Y$. The marginal bpa on X, $m^{\downarrow X}$ (and similarly on Y, $m^{\downarrow Y}$), is defined in the following way:

$$m^{\downarrow X}(A) = \sum_{R|A=R_X} m(R), \quad \forall A \subseteq X \tag{1}$$

where R_X is the set projection of R on X.

There are two functions associated with each basic probability assignment, a belief function, Bel, and a plausibility function, Pl: $Bel(A) = \sum_{B \subseteq A} m(B), \quad Pl(A) = \sum_{A \cap B \neq \emptyset} m(B)$. They can be seen as the lower and upper probability of A, respectively.

We may note that belief and plausibility are interrelated for all $A \in \wp(X)$, by $Pl(A) = 1 - Bel(A^c)$, where A^c denotes the complement of A. Furthermore, $Bel(A) \leq Pl(A)$.

On each bpa on a finite set X, there exists a set of associated probability distributions p on X, of the following way:

$$K_m = \{p|\ Bel(A) \leq p(A),\ \forall A \in \wp(X)\} \tag{2}$$

We remark that $Bel(A) \leq p(A)$ is, in this case, equivalent to $Bel(A) \leq p(A) \leq Pl(A)$. K_m is a closed and convex set of probability distributions, also called *credal set* in the literature.

3. Measures of uncertainty in DST

The classical measure of entropy [22] on probability theory is defined by the following continuous function:

$$S(p) = -\sum_{x \in X} p(x) \log_2(p(x)), \tag{3}$$

where $p = (p(x))_{x \in X}$ is a probability distribution on X, $p(x)$ is the probability of value x and $\log_2$ is normally used to quantify the value in bits.[1] The value $S(p)$ quantifies the only type of uncertainty presented on probability theory and it verifies a large set of desirable properties [18,22].

In DST, Yager [23] makes the distinction between two types of uncertainty. One is associated with cases where the information is focused on sets with empty intersections and the other one is associated with cases where the information is focused on sets with cardinality greater than one. They are called *discord* (*randomness* or *conflict*); and *non-specificity* respectively. So far, both types of uncertainty have been considered with the same level of importance in DST.

The following function, introduced by Dubois and Prade [9], has its origin in the classical Hartley measure [10] on classical set theory, and on the extended Hartley measure on possibility theory (Higashi and Klir [13]). It represents a measure of non-specificity associated with a bpa and it is expressed as follows:

$$I(m) = \sum_{A \subseteq X} m(A) \log_2(|A|). \tag{4}$$

$I(m)$ attains its minimum, zero, when m is a probability distribution. The maximum, $\log_2(|X|)$, is obtained for a bpa, m, with $m(X) = 1$ and $m(A) = 0,\ \forall A \subset X$. It is showed in the literature that I verifies all the required properties for such a type of measure. It was extended on more general theories than DST in Abellán and Moral [3].

Many measures were introduced to quantify the discord degree that a bpa represents [18]. One of the most representative discord functions was introduced by Yager [23]:

$$E(m) = -\sum_{A \subseteq X} m(A) \log_2 Pl(A). \tag{5}$$

But this function does not verify in DST all the required properties.

We can see in the literature, about measures of uncertainty in DST, that when it is presented a new composed measure, i.e. a measure which quantifies both types of uncertainty (a TU measure), then both parts of uncertainty are considered with the same weight. We mean that the upper and the lower possible values for each part are the same. We can cite some examples of this type of composite measures with the same weight for both parts: Lamata and Moral [19], Klir and Ramer [17], Klir and Parvitz [16], Maeda and Ichihashi [20], Abellán, Klir and Moral [1]. On the other hand, we think that it could be discussed if we consider that the non-specificity part can suppose an important difference between the DST and the PT, where only the part of discord appears.

Harmanec and Klir [11,12] proposed the measure $S^*(m)$, equal to the maximum of the entropy (upper entropy) of the probability distributions verifying $Bel(A) \leq \Sigma_{x \in A} p(x) \leq Pl(A),\ \forall A \subseteq X$. This set of probability distributions is the credal set associated with a bpa m, that we have noted as K_m.

Harmanec and Klir [12] proposed S^* as a total uncertainty measure in DST, but they do not separate both parts. Abellán, Klir and Moral [1], have proposed upper entropy as an aggregate measure on more general theories than DST, separating coherently discord and non-specificity. These parts can be also obtained in DST in a similar way. In DST, one can consider

$$S^*(m) = S_*(m) + (S^* - S_*)(m), \tag{6}$$

where $S^*(m)$ represents maximum entropy and $S_*(m)$ represents minimum entropy on the credal set K_m associated to a bpa m, with $S_*(m)$ coherently quantifying the discord part and $(S^* - S_*)(m)$ its non-specificity part. That measure has been successfully used in applications (see Abellán and Moral [4]). An algorithm to obtain that value in DST, and on more general theories, can be found in Abellán and Moral [6].

Very recently, Deng [7] have presented a new uncertainty measure named *Deng entropy* that can be considered as a new composed measure, quantifying discord and non-specificity.

This function, called E_d can be defined as follow for a bpa m on a finite set X:

$$E_d(m) = -\sum_{A \subseteq X} m(A) \log_2 \frac{m(A)}{2^{|A|} - 1}. \tag{7}$$

It can be separated in two functions measuring the both types of uncertainty in DST:

$$E_d(m) = -\sum_{A \subseteq X} m(A) \log_2 m(A) + \sum_{A \subseteq X} m(A) \log_2(2^{|A|} - 1), \tag{8}$$

where the first term quantifies the part of discord and the second one the part of non-specificity of a bpa.

The E_d measure arises with the idea to give more importance to the increase in uncertainty produced when the number of alternatives increases, i.e. on the non-specificity part. It is not agreed with the standard bounds of values for such type of measures. It can be observed that the upper bound for the part of discord can be notably smaller than the one for the non-specificity part. This will be analyzed in the following sections.

The part of discord of E_d is a natural extension of the Shannon's entropy and was studied in the 90′s. It verifies a set of interesting properties but not all the necessary ones. In the literature, it has been considered that there is no discord in a bpa when all the focal sets share, at least, an element (see Abellán and Moral [5], Abellán and Masegosa [2]). The E measure of Yager is also agree with that consideration, and attains a value of 0 in that case. But, in that situation, the measure used in E_d to quantify the discord can express a positive value, that is not coherent with that concept of discord.

The part of non-specificity of E_d is some similar to the Hartley measure but it emphasizes in a very strong way on the number

[1] Indifferently, log and $\log_2$ are used in the literature for this aim.

of possible alternatives, giving us greater values than the ones obtained from the Hartley measure.

3.1. Properties of TU measures in DST

Klir and Wierman [18] defined five requirements for a total uncertainty measure (*TU*) in DST, i.e. for a measure which captures both discord and non-specificity. These requirements are the following ones:

(P1) **Probabilistic consistency**: When all the focal elements of a bpa m are singletons then the total uncertainty measure must be equal to the Shannon entropy:

$$TU(m) = \sum_{x \in X} m(x) \log_2 m(x). \tag{9}$$

(P2) **Set consistency**: When exist a set A such that $m(A) = 1$ then a *TU* measure must collapse to the Hartley measure:

$$TU(m) = \log_2 |A|. \tag{10}$$

(P3) **Range**: The range of a $TU(m)$ measure must be $[0, \log_2|X|]$.

(P4) **Subadditivity**: Let m be a bpa on the space $X \times Y$, $m^{\downarrow X}$ and $m^{\downarrow Y}$ its marginal bpas on X and Y respectively, then a *TU* measure must satisfy the following inequality:

$$TU(m) \le TU(m^{\downarrow X}) + TU(m^{\downarrow Y}). \tag{11}$$

(P5) **Additivity**: Let m be a bpa on the space $X \times Y$, $m^{\downarrow X}$ and $m^{\downarrow Y}$ its marginal bpas on X and Y respectively such that these marginal are not interactive $(m(A \times B) = m^{\downarrow X}(A)\, m^{\downarrow Y}(B)$, with $A \subseteq X$, $B \subseteq Y$ and $m(C) = 0$ if $C \neq A \times B$), then a *TU* measure must satisfy the equality:

$$TU(m) = TU(m^{\downarrow X}) + TU(m^{\downarrow Y}). \tag{12}$$

The set of properties verified by Shannon's entropy in probability theory was the starting point to enunciate these ones in DST. In DST there is more types of uncertainty than in classic PT. Hence, the requirement of range could be debatable. In the literature, one can find arguments in favor of a larger range.

In DST, we can find situations that never appear in probability theory. A probability distribution never contains other probability distribution; but in DST, the information of a bpa can be contained by the information of another bpa, as we can see in the following example.

Example 1. *Situation1*: A policeman has 3 pieces of evidence: Ev_1, Ev_2 and Ev_3; that express the degree of culpability that a thief (T_1, T_2 or T_3) has about a stole. The expert policeman quantifies the information available via a basic probability assignment. He considers the following bpa on the set $X = \{T_1, T_2, T_3\}$:

$$Ev_1 \longrightarrow m_1(\{T_1, T_2, T_3\}) = 0.3,$$
$$Ev_2 \longrightarrow m_1(\{T_1, T_3\}) = 0.2,$$
$$Ev_3 \longrightarrow m_1(\{T_2, T_3\}) = 0.5.$$

Situation2: assume now, that the policeman finds information to discard T_3 in Ev_1 and he needs to modify his evidence to the following ones:

$$Ev_1 \longrightarrow m_2(\{T_1, T_2\}) = 0.3,$$
$$Ev_2 \longrightarrow m_2(\{T_1, T_3\}) = 0.2,$$
$$Ev_3 \longrightarrow m_2(\{T_2, T_3\}) = 0.5.$$

Situation1 represents a situation with a greater level of uncertainty than *Situation2* (a new information has appeared). Here, we have that $Bel_1(A) \le Bel_2(A)$ and $Pl_2(A) \le Pl_1(A)$, $\forall A \subseteq X$, implying a bigger level of uncertainty for m_1.

The situation expressed by Example 1 must be taken into account for a total uncertainty measure in DST and consequently, the following property is required (Abellán and Masegosa [2]):

(P6) **Monotonicity**: A *TU* measure in DST must take into account the decreasing or increasing in information.
Formally, let 2 bpas be on a finite set X, m_1 and m_2, verifying that $Bel_2(A) \le Bel_1(A)$, $\forall A \subseteq X$,[2] then it must be verified that:

$$TU(m_1) \le TU(m_2). \tag{13}$$

From Klir and Wierman [18] and Abellán and Masegosa [2] we know that the upper entropy function verifies the properties P1, P2, P3, P4, P5 and P6. So far, S^* is the only *TU* measure that satisfies all the proposed properties.

4. Properties of Deng entropy

We are going to show what is the set of basic properties verified by E_d.

P1: As it is presented in Deng [7], the E_d measure expresses the same value than the one of Shannon's entropy when it is applied in PT. Hence the P1 property is verified.

P2: The set consistency is broken due to the definition of E_d. It does not coincide with the Hartley measure. When exists $A \subset X$ such that $m(A) = 1$, then

$$E_d(m) = \log_2(2^{|A|} - 1) > \log_2(|A|),$$

when $|A| > 1$ (in other situation we have no uncertainty). As we said, the part of non-specificity of E_d can be considered as an increment of the Hartley measure, but by its definition it does not verify similar properties.

P3: About the range, we think that it is very questionable that for a measure in a theory with 2 types of uncertainty: discord and non-specificity types, the range be the same than the one in a theory with only uncertainty of type discord.
In the Example 4.3 from Deng [7], it is said that the maximum value for E_d is attained in the following mass assignment:[3]

$$m^*(A_i) = \frac{2^{|A_i|} - 1}{\sum_i 2^{|A_i|} - 1}, \quad \forall A_i \subseteq X$$

The maximum value for E_d can be obtained replacing this bpa in the expression of the E_d:

$$E_d(m^*) = -\sum_{A \subseteq X} \frac{2^{|A|} - 1}{\sum_{B \subseteq X} 2^{|B|} - 1} \log_2 \frac{\frac{2^{|A|}-1}{\sum_{B \subseteq X} 2^{|B|}-1}}{2^{|A|} - 1}$$
$$= -\frac{1}{\sum_{B \subseteq X} 2^{|B|} - 1} \sum_{A \subseteq X} \left(2^{|A|} - 1\right) \log_2 \frac{1}{\sum_{B \subseteq X} \left(2^{|B|} - 1\right)}$$
$$= \log_2 \sum_{B \subseteq X} \left(2^{|B|} - 1\right)$$

Hence, the range of E_d is $[0, \log_2 \sum_{B \subseteq X} \left(2^{|B|} - 1\right)]$, that is clearly greater than $[0, \log_2(n)]$, with $n = |X|$.

P4: To check that E_d does not verify the subadditivity property, we consider the following example:

Example 2. Let $X \times Y$ be the product space of the sets $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2\}$, and m on $X \times Y$ with masses

$$m(\{z_{11}, z_{12}, z_{21}\}) = 0.6, \quad m(\{z_{31}, z_{32}\}) = 0.1,$$
$$m(X \times Y) = 0.3,$$

[2] It is called *Weak Monotonicity*.

[3] It is supposed that in Deng [7] that mass assignment has been obtained maximizing the expression of E_d with respect to the values of $m(A)$, but this is not said.

where $z_{ij} = (x_i, y_j)$. We have that the marginal bpas on X and Y are the following ones: m^1 and m^2, respectively:

$$m^1(\{x_1, x_2\}) = 0.6, \quad m^1(\{x_3\}) = 0.1, \quad m^1(X) = 0.3;$$
$$m^2(Y) = 1.$$

Now, the values of uncertainty via E_d are the following ones:

$$E_d(m) = 1.2954 + 3.6361 = 4.9315;$$
$$E_d(m^1) + E_d(m^2) = 1.2954 + 1.7932 + 0 + 1.5850 = 4.6736.$$

Hence, $E_d(m) > E_d(m^1) + E_d(m^2)$ and the subadditivity property is not satisfied.

P5: E_d is also non additive. It is easy to check it because the discord part is additive and the non-specificity part is not additive because, in general, $2^{nm} - 1 \neq (2^n - 1)(2^m - 1)$. Then the sum of both parts is not additive. We can use the following counterexample to prove it in a more direct way:

Example 3. With the notation of the Example 2, let $X \times Y$ be the product space of the sets $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2\}$, and m^1 and m^2 the following bpas on X and Y, respectively:

$$m^1(\{x_1, x_2\}) = 0.6, m^1(\{x_3\}) = 0.1, m^1(X) = 0.3;$$
$$m^2(Y) = 1.$$

Now, we build the following bpa $m' = m^1 \times m^2$ on $X \times Y$ (the marginal bpas of m' are m^1 and m^2; and they are non interactive). The bpa m' has the following masses:

$$m'(\{z_{11}, z_{12}, z_{21}, z_{22}\}) = 0.6, \quad m'(\{z_{31}, z_{32}\}) = 0.1,$$
$$m(X \times Y) = 0.3,$$

where $z_{ij} = (x_i, y_j)$.
The values of uncertainty via E_d are the following ones:

$$E_d(m') = 1.2954 + 4.2958 = 5.5912;$$
$$E_d(m^1) + E_d(m^2) = 1.2954 + 1.7932 + 0 + 1.5850 = 4.6736.$$

Again $E_d(m') > E_d(m^1) + E_d(m^2)$, and the additivity property is not satisfied by E_d.

We can resume that the E_d measure only verifies the P1 property of the first 5 basic properties, i.e. of the properties proposed in Klir and Wierman [18].

We consider that the P6 property about the monotonicity is an important property for TU measures in DST, and it can be considered as an essential property. A non-monotonous TU can give us incoherent results when we need to choose the most informative bpa in applications. Hence we study it separately. A clear decreasing or increasing in information must not be reflected via an inverse way by a TU measure in DST, it would have no sense. In the following example we see that E_d measure has a bad behavior in that line.

Example 4. Let m^1, m^2 be the following bpas on the space $X = \{x_1, x_2\}$, with masses

$$m^1(X) = 1;$$
$$m^2(\{x_1\}) = t, \quad m^2(X) = 1 - t.$$

We see that, when $t > 0$, m^1 expresses a situation of less information than m^2 because m^2 points to $\{x_1\}$. With that information (m^2), if we must bet on one alternative, we choose $\{x_1\}$ because the evidence supports that alternative in a stronger way. Also we have a situation of weak monotonicity between these two bpas: $Bel_1(A) \leq Bel_2(A), \forall A \subseteq X$. But, the values of uncertainty via E_d are the following ones:

$$E_d(m^1) = \log_2(3);$$
$$E_d(m^2) = -t\log_2(t) - (1-t)\log_2(1-t) + t\log_2(3).$$

Considering $E_d(m^2)$ as a function of $t \in [0, 1]$, it is easy to calculate that it has its maximum value when $t = 0.429$. It can be checked that the intervals of belief ($[Bel, Pl]$) from m^2 are very different for both alternatives: $[0.429, 1]$ for $\{x_1\}$; and $[0, 0.571]$ for $\{x_2\}$. We clearly would bet on the first alternative. However, with the m^1 bpa, we have no information on which alternative we must bet.

In this case we have that $E_d(m^1) < E_d(m^2)$:

$$E_d(m^1) = 1.58 < E_d(m^2) = 1.665,$$

representing that m^1 expresses a situation of more information (less uncertainty) than m^2, but this is not intuitively correct.

With the above example, we see that E_d also does not verify the Monotonicity property. It represents a situation of incoherent values taking into account the amount of information expressed for each bpa.

We must expect that, in situations as the one exposed in Example 4, where a bpa m^1 represents more uncertainty than other bpa m^2, a *TU* measure in DST "at least" must give us a value for m^2 that *be not greater than* the one for m^1.

4.1. Some questionable behaviors of Deng entropy

We are going to present some questionable performing of the Deng entropy due to its definition. We find that the behavior of this new measure breaks something with the classical behavior of the TU measures presented as far. Under our point of view, its behavior has some sense in determinate situations.

First of all, as a direct consequence of its definition, Deng entropy distinguishes between "no information" and "equal certainty", that it is very questionable. The first one is expresses, for example, by a bpa m^1 on a finite set $X = \{x_1, x_2, x_3\}$, such that $m^1(X) = 1$; and the second one by a bpa m^2, where $m^2(x_j) = 1/3$, for all j. As we can see in Deng [7], the following values are obtained:[4]

$$E_d(m^2) = 1.5850 < E_d(m^1) = 2.8074$$

If we consider the real information expressed by both evidences, we could consider that we have the same information (uncertainty). We use the following Example 5 to better understand that assertion.

Example 5. Consider a case about a bet about a race among 3 cars. We have two experts that have the following information:

Expert1: "The 3 cars and pilots are very very similar"
Expert2: "I do not have information about the characteristics of each car and pilot"

The opinion of the *Expert1* produces the uniform probability distribution (1/3, 1/3, 1/3); and the second one the vacuous basic probability assignment (above m^1 and m^2 bpas). The question is if we must to bet on a car, on which one we must to do it following the information of each expert?. We do not know on which one to bet, we have no information to know something that allow us to bet on a determinate car. Hence, the information (uncertainty) should be the same. In both cases, it must be the maximum possible uncertainty value.

[4] Using $\log_2$ function.

That reasoning in Example 5 is compatible with the *Maximum Uncertainty Principle* in DST (Klir and Wierman [18], Klir [15]). This principle stands on that our total ignorance (lack of knowledge) must be always recognized. It comes from the *Maximum Entropy Principle* from Jaynes [14] who expressed the following sentence:

"... maximum-entropy principle that it achieves 'objectivity' of our inferences in the sense that we base our predictions only on the information that we do, in fact, have... ".

A second debatable characteristic of this measure is the difference between the two parts used to quantify the discord and the non-specificity. The first one has a range of $[0, \log_2(n)]$, with n the cardinal of the whole set; and the one of the second part is $[0, \log_2(2^n - 1)]$. That difference increases when the number of alternatives increases. In those cases the possible values from the discord part can be notably smaller than the ones from the non-specificity part. For example, when $n = 10$ the maximum possible value for the discord part is 3.3 and the one for the non-specificity part is 10; but increasing n to 20, we have values of 4.3 and 20, respectively. In this line the discord part can have few importance in a bpa when the number of possible alternatives increases. In our opinion, that is not a consideration without sense, but it is some questionable and breaks with the thoughts from 90′s that the both part of uncertainty in DST have the same weight, as it was expressed in Section 2.

A third drawback found on the definition of the E_d was expressed in previous sections and it is about its discord part. This part can be positive in situations where there is no discord, if we consider that the discord does not appear when all the focal set share an element.[5] For example, for the bpa m on $X = \{x_1, x_2, x_3\}$, such that $m(\{x_1\}) = m(X) = 0.5$, we have no discord but the discord part of E_d does not express that:

$$-\sum_{A\subseteq X} m(A)\log_2 m(A) = 1.$$

5. Conclusions and future works

This paper studied the properties of a new measure of total uncertainty in DST called Deng entropy. We used as a reference the set of properties well accepted in the literature for this type of measures, and other new one about the monotonicity. We shown that the Deng entropy verifies only one of those properties. Some of the basic properties could be discussed but other ones are essential, intuitively correct and necessary for a measure of that type; and unfortunately are not satisfied by the Deng entropy. Other debatable behaviors of that measure have been exposed.

Along the paper it has been remarked some considerations about the range of a *TU* measure in DST; and about the possible different importance of each part of uncertainty in DST. We consider them as important tasks that need more studies. With respect to the property of the range, we think that it should be enlarged, because in DST we have more types of uncertainty than in PT. Also, considering the Example 3.1 in the paper of Deng [7], one can think that the non-specifity part can have more weight than the discord part in DST. We will consider those aspects for future works.

A continuity property could be also considered, taking into account that an uncertainty measure must give us similar values when it is applied on similar evidences. But it is necessary to take into account the appropriate distance functions in DST to express it in a formal way. It can represent an interesting task for our future work.

Considering all the results in this paper, we think that the Deng entropy should be used in applications only in a very cautious way due its shortcomings, principally motivated by the no verification of important properties.

Acknowledgments

This work has been supported by the Spanish "Ministerio de Economía y Competitividad" and by "Fondo Europeo de Desarrollo Regional" (FEDER) under Project TEC2015-69496-R.

References

[1] Abellán J, Klir GJ, Moral S. Disaggregated total uncertainty measure for credal sets. Int J Gen Syst 2006;35(1):29–44.
[2] Abellán J, Masegosa A. Requierements for total uncertainty measures in dempster-shafer theory of evidence. Int J Gen Syst 2008;37(6):733–47.
[3] Abellán J, Moral S. A non-specificity measure for convex sets of probability distributions. Int J Uncertainty Fuzziness Knowl Based Syst 2000;8:357–67.
[4] Abellán J, Moral S. Upper entropy of credal sets. applications to credal classification. Int J Approximate Reasoning 2005;39(2–3):235–55.
[5] Abellán J, Moral S. Difference of entropies as a non-specificty measure for credal sets. Int J Gen Syst 2005;34(3):201–14.
[6] Abellán J, Moral S. An algorithm that computes the upper entropy for order-2 capacities. Int J Uncertainty Fuzziness Knowl Based Syst 2005;14(2):141–54.
[7] Deng Y. Deng entropy. Chaos, Solitons Fractals 2016;91:549–53.
[8] Dempster AP. Upper and lower probabilities induced by a multivalued mapping. Ann Math Stat 1967;38:325–39.
[9] Dubois D, Prade H. A note on measure of specificity for fuzzy sets. BUSEFAL 1984;19:83–9.
[10] Hartley RVL. Transmission of information. Bell Syst Tech J 1928;7:535–63.
[11] Harmanec D, Klir GJ. Measuring total uncertainty in dempster-shafer theory: a novel approach. Int J Gen Syst 1994;22:405–19.
[12] Harmanec D, Klir GJ. Principle of uncertainty revisited. proc. 4th intern. Fuzzy Syst Intell Control Conf, Maui, Hawai 1996:331–9.
[13] Higashi M, Klir GJ. Measures of uncertainty and information based on possibility distributions. Int J General Syst 1983;9:43–58.
[14] Jaynes ET. Where do we stand on maximum entropy?. Levine RD, Tribus M, editors MIT Press: Cambridge, MA, USA; 1979.
[15] Klir GJ. Principles of uncertainty: what are they?. why do we need them? Fuzzy Sets Syst 1995;74:15–31.
[16] Klir GJ, Parviz B. A note on the measure of discord, proceedings of the eighth conference on uncertainty in artificial intelligence. Stanford, California, USA; 1992. p. 138–41.
[17] Klir GJ, Ramer A. Uncertainty in the dempster-shafer theory: a critical reexamination. Int J Gen Syst 1990;18(2):155–66.
[18] Klir GJ, Wierman MJ. Uncertainty-Based information. Phisica-Verlag; 1998.
[19] Lamata M, Moral S. Measures of entropy in the theory of evidence. Int J Gen Syst 1987;14:297–305.
[20] Maeda Y, Ichihashi H. A uncertainty measure with monotonicity under the random set inclusion. Int J Gen Syst 1993;21:379–92.
[21] Shafer G. A mathematical theory of evidence. Princeton University Press, Princeton; 1976.
[22] Shannon CE. A mathematical theory of communication. Bell Syst Tech J 1948;27:379–423. 623–656
[23] Yager RR. Entropy and specificity in a mathematical theory of evidence. Int J Gen Syst 1983;9:249–60.

[5] We remember that the discord uncertainty appears when the bpa has focal sets with no common elements.