Chaos, Solitons and Fractals 140 (2020) 110112

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

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

Critique of modified Deng entropies under the evidence theory

Serafín Moral-García*, Joaquín Abellán

Department of Computer Science and Artificial Intelligence, University of Granada, Granada, Spain

ARTICLE INFO

Article history:
Received 1 April 2020
Revised 3 July 2020
Accepted 9 July 2020
Available online 14 July 2020

Keywords:
Evidence theory
Uncertainty measures
Deng entropy
Modified Deng entropies
Conflict
Non-specificity

ABSTRACT

The Evidence theory or Dempster-Shafer Theory (DST) has been frequently used in practical applications to deal with uncertainty or lack of information. It is based on the concept of basic probability assignment (BPA). In DST, it is important to quantify the uncertainty (or information) that a BPA represents. An uncertainty measure, known as Deng entropy, was introduced as an interesting alternative to other measures proposed before. In previous work, it was shown that the Deng entropy does not verify most of the required properties for this type of measure and presents some undesirable behaviors. Two modifications of the Deng entropy have been recently proposed, which improve the original one. In this research, we demonstrate that these modifications do also not satisfy the majority of the necessary mathematical properties, and they present most of the behavioral drawbacks of the original one. Therefore, as the original Deng entropy, the modified ones should be cautiously employed in practical applications.

1. Introduction

In many situations, the use of the Probability Theory (PT) to represent the information is not very suitable because the available information is not reliable enough. For this reason, many mathematical theories based on imprecise probabilities have been developed in the literature [1]. All of them generalize PT.

The *Evidence theory* or *Dempster Shafer theory* (DST) [2,3] has been frequently used in the literature to deal with uncertainty in several domains such as *medical diagnosis* [4], *statistical classification* [5], *face recognition* [6], or *target identification* [7]. DST is based on the *basic probability assignment* concept (BPA), a generalization of the concept of probability distribution in PT.

In DST, it is quite important to quantify the uncertainty-based information that a BPA represents. For this purpose, many approaches have been proposed so far. Most of the uncertainty measures developed in DST take as a reference Shannon's entropy [8] for probability distributions, which satisfies a large set of desired properties.

Since DST includes PT, in DST, more types of uncertainty can be represented than in PT. Yager, in [9], distinguishes between two types of uncertainty in DST. The first one of them is called *conflict*, which appears when the information is focused on sets with empty intersection. The second type is called *non-specificity*, which corresponds to cases where the information is focused on sets with cardinality greater than one. The difference between uncertainty in PT and DST principally resides in the non-specificity part.

In [10], it was exposed a study about the set of mathematical requirements that a measure that quantifies both types of uncertainty present in DST must satisfy. That study was extended in [11]. Some of these requirements are pretty questionable, but other ones are crucial. For example, if a BPA is defined over a finite set that can be decomposed on two more simple ones, the total amount of uncertainty must not be decreased with such a decomposition. Furthermore, if we join two non-interactive BPAs, the total amount of information must not vary. Also, a total uncertainty measure has to take into consideration coherently an increase or decrease of information. So far, the upper entropy, presented in [12], is the unique uncertainty measure in DST that verifies all the requirements [11].

However, the algorithms for the calculation of the upper entropy, proposed in [12–14], are notably complex. For this reason, an interesting total uncertainty measure, called *Deng entropy*, was proposed in [15]. According to this measure, the amount of uncertainty is strongly influenced by the number of possible alternatives. As shown in [16], this measure presents several drawbacks and, thus, we should be cautious when we use it in practical applications; it does not verify most of the necessary mathematical properties, and its behavior in many situations is undesirable.

In [17], it was proposed a modification of the Deng entropy that presents improvement since it takes into account the total number of possible alternatives to a higher degree. Another version of the Deng entropy was proposed in [18]. It performs better than the original Deng entropy because it considers the intersection between statements on uncertainty.

* Corresponding author.
E-mail address: seramoral@decsai.ugr.es (S. Moral-García).

https://doi.org/10.1016/j.chaos.2020.110112

In this work, we demonstrate that these modifications of the Deng entropy do also not verify most of the essential mathematical properties such as subadditivity, additivity, or monotonicity. Moreover, we also show that, as the original Deng entropy, the behavior of the modified ones in some situations is also questionable. Therefore, the modified Deng entropies should be cautiously utilized in practical applications, as the original one.

This paper is arranged as follows: Section 2 describes the Dempster-Shafer Theory, the most important uncertainty measures proposed in DST, and the set of mathematical properties that an uncertainty measure in DST should satisfy. The properties verified by the modified Deng entropies are exposed in Section 3. Section 4 presents some questionable behaviors of the modifications of the Deng entropy considered here. Conclusions and ideas for future research are given in Section 5.

2. Background

2.1. Dempster-Shafer theory of evidence

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite set considered as a set of possible situations. Let $\wp(X)$ be the set of all the subsets in X.

The *Evidence theory*, also known as *Dempster Shafer theory* (DST) [2,3], is based on the concept of *basic probability assignment* (BPA). It consists of a mapping m: $\wp(X) \to [0, 1]$ such that $\sum_{A\in\wp(X)} m(A) = 1$, and $m(\emptyset) = 0$.

If a subset $A \subseteq X$ verifies that $m(A) > 0$ it is said that “A is a *focal element* of m”.

Two functions are associated with each BPA: a *belief function* and a *plausibility function*. They are defined as follows:

$$Bel(A) = \sum_{B|B\subseteq A} m(B), \quad Pl(A) = \sum_{B|B\cap A\neq\emptyset} m(B), \quad \forall A \in \wp(X). \tag{1}$$

Obviously, $Bel(A) \le Pl(A)$ $\forall A \in \wp(X)$. In addition:

$$Pl(A) = 1 - Bel(\bar{A}), \quad \forall A \in \wp(X). \tag{2}$$

being $\bar{A}$ the complementary of A.

For each BPA m on X, there exists a closed and convex set of probability distributions (also called credal set) associated with it. Such a set is defined in the following way:

$$\mathcal{P}_m = \{p \in \mathcal{P}(X) \mid Bel(A) \le p(A), \quad \forall A \in \wp(X)\}, \tag{3}$$

where $\mathcal{P}(X)$ is the set of all probability distributions on X.

Let us suppose now that X and Y are finite sets, and m is a BPA on $X \times Y$. The marginal BPA on X, denoted as $m^{\downarrow X}$, is defined in the following way:

$$m^{\downarrow X}(A) = \sum_{R|A=R_X} m(R), \quad \forall A \in \wp(X), \tag{4}$$

being R_X the projection of R on X:

$$R_X = \{x \in X \mid \exists y \in Y, (x, y) \in R\}, \quad \forall R \subseteq X \times Y. \tag{5}$$

The definition of the marginal BPA on Y, $m^{\downarrow Y}$, is analogous.

2.2. Uncertainty measures in DST

It is well-established that uncertainty in classical Probability Theory (PT) is measured via the Shannon entropy [8], which is defined as follows:

$$S(p) = \sum_{x\in X} p(x)\log_2(p(x)), \tag{6}$$

being $p = (p(x))_{x\in X}$ a probability distribution on X. The function S captures the type of uncertainty called *conflict*, the only one existing in PT. It satisfies a set of desirable properties [8,10].

On the other hand, the Hartley measure [19] is known to be suitable to quantify uncertainty in classical possibility theory. It is defined by:

$$H(A) = \log_2(|A|), \quad \forall A \subseteq X. \tag{7}$$

The type of uncertainty measured by H is different from the one quantified by the Shannon entropy; it is usually called *non-specificity*.

Yager [9] establishes that, in DST, both types of uncertainty: conflict and non-specificity, coexist. In DST, conflict corresponds to cases where the information is focused on disjunct sets; and non-specificity appears when the information is focused on sets whose cardinality is greater than one.

In [20], it was proposed a generalization of the Hartey measure to DST, defined in the following way:

$$GH(m) = \sum_{A\in\wp(X)} m(A)\log_2(|A|). \tag{8}$$

When m is a probability distribution, GH attains its minimum value, which is equal to 0. Its maximum value, $\log_2(|X|)$, is obtained when $m(X) = 1$. GH is a well-established non-specificity measure in DST; it satisfies a set of desirable properties. Furthermore, it can be extended to more general theories than DST [21].

In order to measure the conflict in a BPA, many measures have been presented in the literature. One of the most remarkable of them was introduced in [9]. It is defined in the following way:

$$E(m) = -\sum_{A\in\wp(X)} m(A)\log_2 Pl(A). \tag{9}$$

Nevertheless, this function does not satisfy all the crucial properties for uncertainty measures in DST.

In the literature, when it is presented a measure that jointly quantifies conflict and non-specificity, both parts often have the same weight. Examples of this point can be found in [22–25]. Nonetheless, this issue could be discussed if we think that the non-specificity part could have a higher weight because it is the main difference between the uncertainty in DST and PT.

Harmanec and Klir, in [12], proposed a measure of total uncertainty in DST. It consists of the maximum of entropy $S^*(m)$ among the probability distributions belonging to the credal set associated with m, $\mathcal{P}_m$, defined in Eq. (3). Nevertheless, they do not separate non-specificity and conflict.

Abellán, et al., in [22], proposed S^* as an aggregate measure that coherently separates non-specificity and conflict on more general theories than DST. This separation is also valid for DST. We can consider:

$$S^*(m) = S_*(m) + (S^* - S_*)(m), \tag{10}$$

being $S_*(m)$ the minimum of entropy on $\mathcal{P}_m$. $S_*(m)$ captures conflict and $(S^* - S_*)(m)$ indicates non-specificity. This measure has been successfully employed in practical applications [26]. Algorithms to calculate S^* can be found in [12–14].

However, these algorithms are notably complex. For this reason, a total uncertainty measure known as *Deng entropy* was presented in [15]. It also jointly quantifies non-specificity and conflict. It is defined as follows:

$$E_d(m) = -\sum_{A\in\wp(X)} m(A)\log_2\left(\frac{m(A)}{2^{|A|}-1}\right). \tag{11}$$

This function can be re-written in the following way [15]:

$$E_d(m) = \sum_{A\in\wp(X)} m(A)\log_2\left(2^{|A|}-1\right) - \sum_{A\in\wp(X)} m(A)\log_2 m(A). \tag{12}$$

In the above formula, the first term measures the non-specificity part, and the second one quantifies the conflict part. The idea of this measure is that the uncertainty must be considerably increased as there are more alternatives.

More recently, in [17], it was proposed a modification of the Deng entropy, known as Zhou entropy. It is defined in the following way:

$$E_{Zhou}(m) = -\sum_{A\in\wp(X)} m(A)\log_2\left(\frac{m(A)}{2^{|A|}-1}\exp\left(\frac{|A|-1}{|X|}\right)\right). \quad (13)$$

Similarly as happens with the Deng entropy, this function can be re-written as follows:

$$\begin{aligned} E_{Zhou}(m) = &\sum_{A\in\wp(X)} m(A)\log_2(2^{|A|}-1) \\ &-\sum_{A\in\wp(X)} m(A)\log_2\left(\exp\left(\frac{|A|-1}{|X|}\right)\right) \\ &-\sum_{A\in\wp(X)} m(A)\log_2 m(A). \end{aligned} \quad (14)$$

It could be considered that the first two terms in Eq. (14) quantify the non-specificity part in a BPA since both of them are equal to 0 when m is a probability distribution. The third one might measure the conflict part, which is the same as in the original Deng entropy. As it can be observed, E_{Zhou} is also based on the idea of the Deng entropy as it gives a higher total uncertainty value when the number of alternatives increases. Nonetheless, due to the second term, with this modification, this increase is more controlled.

Afterward, in [18], Cui, Liu, Zhang, and Kang proposed a new version of the Deng entropy that takes into consideration the intersections between the focal sets. It is defined in the following way:

$$\begin{aligned} E_{Cui}(m) = &-\sum_{A\in\wp(X)} m(A)\log_2 \\ &\times\left[\left(\frac{m(A)}{2^{|A|}-1}\right)\times\exp\left(\sum_{B\neq A\wedge m(B)>0}\frac{|A\cap B|}{2^{|X|-1}}\right)\right]. \end{aligned} \quad (15)$$

It is possible to re-write this function as follows:

$$\begin{aligned} E_{Cui}(m) = &\sum_{A\in\wp(X)} m(A)\log_2\left(2^{|A|}-1\right) - \sum_{A\in\wp(X)} m(A)\log_2 \\ &\times\left[m(A)\times\exp\left(\sum_{B\neq A\wedge m(B)>0}\frac{|A\cap B|}{2^{|X|-1}}\right)\right]. \end{aligned} \quad (16)$$

In the above expression, the first term indicates non-specificity, while the second one captures conflict. Indeed, when m is a probability distribution, the first term is equal to 0, and the second one collapses to the Shannon entropy. Hence, in Cui entropy, the non-specificity part is the same as in the Deng entropy. However, the conflict part of E_{Cui} is lower than the one of E_d due to the exponential term.

2.3. Required properties of uncertainty measures in DST

Klir and Wierman, in [10], established five essential requirements that every total uncertainty measure (TU) in DST, which jointly captures non-specificity and conflict, must satisfy. These properties are the following ones:

- (P1) **Probabilistic consistency:** If a BPA m is a probability distribution, then a TU measure has to coincide with the Shannon entropy:

$$TU(m) = \sum_{x\in X} m(\{x\})\log_2(m(\{x\})). \quad (17)$$

- (P2) **Set consistency:** If there is a subset $A\subseteq X$ such that $m(A) = 1$, then a TU measure must be equal to the Hartley measure:

$$TU(m) = \log_2|A|. \quad (18)$$

- (P3) **Range:** The range of every TU measure has to be equal to $[0, \log_2|X|]$.
- (P4) **Subadditivity**: Let us consider a BPA m on a product space $X\times Y$. Let $m^{\downarrow X}$ and $m^{\downarrow Y}$ be its marginal BPAs on X and Y, respectively. Then, a TU measure must verify that:

$$TU(m) \le TU(m^{\downarrow X}) + TU(m^{\downarrow Y}). \quad (19)$$

- (P5) **Additivity**: Let us suppose that m is a BPA on a product space $X\times Y$. Let $m^{\downarrow X}$ and $m^{\downarrow Y}$ be its marginal BPAs on X and Y, respectively. Let us assume that the marginals are not interactive, i.e $m(A\times B) = m^{\downarrow X}(A)m^{\downarrow Y}(B)$, $\forall A\subseteq X, B\subseteq Y$, and $m(C) = 0$ if $C\neq A\times B$. Then, for every TU measure, it must be satisfied that:

$$TU(m) = TU(m^{\downarrow X}) + TU(m^{\downarrow Y}). \quad (20)$$

Since in DST there are more types of uncertainty than in PT, the range property, according to which the range of a TU measure in DST must coincide with the one of a TU measure in PT, is debatable. In the literature, arguments favorable to a larger range have been provided.

DST is more general than PT. For this reason, in DST, it can appear situations that never happen in PT. A probability distribution never contains another one. However, the information represented by a BPA in DST can contain the information of another one [11,16]. This issue has to be taken into consideration by uncertainty measures in DST. Therefore, the following property is necessary [11]:

- (P6) **Monotonicity**: In DST, a TU measure has to take into consideration an increase or decrease of information consistently. Formally, if m_1 and m_2 are 2 BPAs on X such that $\mathcal{P}_{m_1}\subseteq\mathcal{P}_{m_2}$, then the following inequality must be satisfied for every TU measure:

$$TU(m_1) \le TU(m_2). \quad (21)$$

The maximum of entropy S^* is the only TU measure in DST proposed so far that satisfies P1-P6 [11].

3. Properties of the modified Deng entropies

We show below which of the required mathematical properties exposed in Section 2.3 are satisfied by the modified Deng entropies E_{Zhou} and E_{Cui}.

- P1: It is easy to observe that, if m is a probability distribution, then the first two terms of Eq. (14) are equal to 0, and E_{Zhou} collapses to the Shannon entropy. In these situations, the first them of Eq. (16) is equal to 0. The same happens with $\sum_{B\subseteq X, B\neq A\wedge m(B)>0}\frac{|A\cap B|}{2^{|X|-1}}$, $\forall A\subseteq X$. Thus, E_{Cui} also coincides with the Shannon entropy for probability distributions. Consequently, both modifications of the Deng entropy satisfy the probabilistic consistency.
- P2: If $m(A) = 1$ for some $A\subseteq X$, then:

$$E_{Zhou}(m) = \log_2(2^{|A|}-1) - \log_2\left(\exp\left(\frac{|A|-1}{|X|}\right)\right).$$

The following result shows that, in this case, if the cardinality of the set of possible alternatives (X) is greater or equal than 3, the value of $E_{Zhou}(m)$ is strictly greater than the one obtained by the generalized Hartley measure.[1]

[1] Except for when A is a singleton, but, in that case, there is no uncertainty.

Proposition 1. *If $|X| \geq 3$ and $m(A) = 1$ for some $A \subseteq X$, then*

$$\log_2 (2^{|A|} - 1) - \log_2 \left(\exp \left(\frac{|A| - 1}{|X|} \right) \right) > \log_2 (|A|),$$

$\forall A \subseteq X,\ |A| \geq 2.$

Proof. It is easy to check that:

$$\log_2 (2^{|A|} - 1) - \log_2 \left(\exp \left(\frac{|A| - 1}{|X|} \right) \right) > \log_2 (|A|) \Leftrightarrow$$

$$\frac{2^{|A|} - 1}{|A|} > \exp \left(\frac{|A| - 1}{|X|} \right).$$

We distinguish 3 cases:

1. $|A| = 2$. In this case:
$$\frac{2^{|A|} - 1}{|A|} = \frac{3}{2} > 1.3956 = \exp \left(\frac{1}{3} \right) \geq \exp \left(\frac{|A| - 1}{|X|} \right).$$
2. $|A| = 3$. Then:
$$\frac{2^{|A|} - 1}{|A|} = \frac{7}{3} > 1.9477 = \exp \left(\frac{2}{3} \right) \geq \exp \left(\frac{|A| - 1}{|X|} \right).$$
3. $|A| \geq 4$. In this case, since the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \frac{2^x - 1}{x}$ is clearly increasing, we have that:
$$\frac{2^{|A|} - 1}{|A|} \geq \frac{2^4 - 1}{4} = \frac{15}{4} > \exp(1) > \exp \left(\frac{|A| - 1}{|X|} \right).$$

□

Therefore, E_{Zhou} does not satisfy the set consistency property. Moreover, it is convenient to remark that there is an unique case where the non-specificity value of E_{Zhou} is lower than the one corresponding to the Hartley measure: $|A| = |X| = 2$. Then,

$$E_{Zhou}(m) = \log_2 (3) - \log_2 \left(\exp \left(\frac{1}{2} \right) \right) = 0.8636 < 1 = \log_2 (|A|).$$

Regarding E_{Cui}, when $m(A) = 1$ for some $A \subseteq X$, $E_{Cui}(m) = \log_2 \left(2^{|A|} - 1\right)$, as the original Deng entropy. What is more, E_{Cui} coincides with E_d when all the focal sets are disjunct, i.e when there is no conflict.
In consequence, E_{Cui} neither satisfies the set consistency property, although, in these cases, unlike E_{Zhou}, E_{Cui} always provides a greater value than the Hartley measure.

- P3:
If $|X| = 4$ and m is a BPA on X such that $m(X) = 1$, then:
$$E_{Zhou}(m) = \log_2 (2^{|X|} - 1) - \log_2 \left(\exp \left(\frac{|X| - 1}{|X|} \right) \right) =$$
$$\log_2 (15) - \log_2 \left(\exp \left(\frac{3}{4} \right) \right) = 2.8249 > 2 = \log_2 (4) = \log_2 (|X|).$$
In such case,
$$E_{Cui}(m) = \log_2 \left(2^{|X|} - 1\right) = \log_2 (15) > \log_2 (4) = \log_2 (|X|).$$
Hence, the range property is not verified by E_{Zhou} nor E_{Cui}.
According to the results proved in [27,28], the maximum value of the Deng entropy is equal to $\log_2 \left(\sum_{A \subseteq X} \left(2^{|A|} - 1\right)\right)$. It is attained with the following BPA:
$$m^*(A) = \frac{2^{|A|} - 1}{\sum_{B \subseteq X} \left(2^{|B|} - 1\right)}, \quad \forall A \subseteq X.$$
It is easy to observe that, in this case, the value obtained by E_{Zhou} is lower than the one attained by E_d due to the second term of Eq. (14). Likewise, in this situation, because of the exponential term of Eq. (15), the value obtained by E_{Cui} is lower than the one achieved by E_d. Consequently, the ranges of E_{Zhou} and E_{Cui} are lower than the range of the original Deng entropy.
- P4: The following example shows that the modified Deng entropies are not subadditive:

Example 1. Let us consider the finite sets $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2\}$. Let m be the following BPA on the product space $X \times Y$:

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

where we have denoted $z_{ij} = (x_i, y_j)$.
The marginal BPAs of m on X and Y, which we denote m^1 and m^2 respectively, are the following ones:

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

$$m^2(Y) = 1.$$

E_{Zhou} takes the following values:

$$E_{Zhou}(m) = 4.2583, \quad E_{Zhou}(m^1) = 2.5116, \quad E_{Zhou}(m^2) = 0.8636$$

We have that $E_{Zhou}(m^1) + E_{Zhou}(m^2) = 3.3752$ and, thus, $E_{Zhou}(m^1) + E_{Zhou}(m^2) < E_{Zhou}(m)$.
Concerning E_{Cui}:

$$E_{Cui}(m) = 4.8674, \quad E_{Cui}(m^1) = 1.4574, \quad E_{Cui}(m^2) = 1.585$$

Hence, $E_{Cui}(m^1) + E_{Cui}(m^2) = 3.0424 < 4.8674 = E_{Cui}(m)$.

- P5: We show in the example below that E_{Zhou} and E_{Cui} do not verify the additivity property. We use the same notation as in Example 1.

Example 2. Let $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2\}$ be finite sets. Let m^1 and m^2 be the following BPAs on X and Y, respectively:

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

$$m^2(Y) = 1.$$

We consider now the following BPA $m' = m^1 \times m^2$ on the product space $X \times Y$. It has the following values:

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

It is easy to check that the marginal BPAs of m' on X and Y are, respectively, m^1 and m^2, and they are non-interactive. We have the following values for E_{Zhou} and E_{Cui}:

$$E_{Zhou}(m') = 4.7873, \quad E_{Zhou}(m^1) = 2.5116, \quad E_{Zhou}(m^2) = 0.8636.$$

$$E_{Cui}(m') = 5.5138, \quad E_{Cui}(m^1) = 1.4574, \quad E_{Cui}(m^2) = 1.585.$$

Thus, $E_{Zhou}(m^1) + E_{Zhou}(m^2) = 3.3752 \neq 4.7873 = E_{Zhou}(m')$, and $E_{Cui}(m^1) + E_{Cui}(m^2) = 3.0424 \neq 5.5138 = E_{Cui}(m')$.

- P6: The following example shows that an increase or decrease of information is not always coherently reflected by E_{Zhou}:

Example 3. Let us consider the finite set $X = \{x_1, x_2\}$ and the following BPAs on X:

$$m^1(X) = 1;$$

$$m^2(\{x_1\}) = m^2(X) = 0.5.$$

We have the following values for E_{Zhou}:

$$E_{Zhou}(m^1) = 0.8636, \quad E_{Zhou}(m^2) = 1.4318.$$

Clearly, the information provided by m^2 is greater than the one expressed via m^1 (m^1 corresponds to total ignorance). Nevertheless,
$E_{Zhou}(m_1) < E_{Zhou}(m_2)$, and E_{Zhou} does not satisfy the monotonicity property.

In the following example, it is shown that the monotonicity requirement is also not verified by E_{Cui}.

Example 4. Let us consider the finite set $X = \{x_1, x_2\}$ and the following BPAs on X:

$$m^1(X) = 0.9, \quad m^1(\{x_1\}) = 0.1;$$

$$m^2(\{x_1\}) = m^2(\{x_2\}) = m^2(X) = \frac{1}{3}.$$

It is easy to check that m^2 expresses more information than m^1. However, $E_{Cui}(m_1) = 1.4146 < 1.4721 = E_{Cui}(m_2)$.

In this way, E_{Zhou} and E_{Cui} only verify the probabilistic consistency property, among the set of mathematical requirements exposed in [10], and extended in [11]. Indeed, some of these properties are debatable, such as the range property (P3). Nonetheless, other properties are crucial: If we have a BPA on a product space, the sum of the uncertainties in the marginal BPAs can not be lower than the uncertainty involved in the original one; if we join two non-interactive BPAs, the total amount of information must not vary; when it is produced an increase of information contained in a BPA, it does not make sense that the uncertainty is increased. The modified Deng entropies considered here present the mentioned shortcomings. The same happens with the original one [16].

4. Some undesirable behaviors of the modified Deng entropies

The original Deng entropy provides incoherent results in some situations because it does not consider the size of the set of possible alternatives correctly [17]. E_{Zhou} was proposed to solve this problem. Also, E_{Cui} improves the original Deng entropy since it considers the intersections between the focal elements. However, as we show in this section, both E_{Zhou} and E_{Cui} also present some behavioral drawbacks, as the original Deng entropy.

- Firstly, the maximum value of E_{Zhou} is not attained with the BPA associated with total ignorance, as we have observed in Example 3, which is an illogical situation because the total ignorance implies a total lack of information. This also happens with the Deng entropy (See [27,28] for more details). In general, since E_{Zhou} does not satisfy the monotonicity property, it is not always consistent with an increase or decrease of information, which is quite undesirable. E_{Cui} neither satisfies the monotonicity property. For this reason, it also obtains incoherent results in some scenarios, as in Example 4.
- As happens with the original Deng entropy, the range of the non-specificity part of E_{Zhou} is greater than the range of the conflict part, although the difference is not as great as with E_d. The difference between both ranges increases as the number of possible alternatives increases. The same happens with E_{Cui}. Thus, the conflict part in both modifications of the Deng entropy might have little importance when there are many alternatives. It could make sense since, as commented before, the main difference between uncertainty in DST and PT resides in the non-specificity part. Nonetheless, it is questionable, and it is not coherent with the thoughts in the literature that both types of uncertainty in DST have the same weight.
- We should remark that, in the original Deng entropy, when the information is focused on one single set, the non-specificity value is always greater than the one corresponding to the Hartley measure. The same happens with E_{Cui}. Indeed, E_d and E_{Cui} obtain identical values when there is a single focal set. When we have two possible alternatives, i.e $X = \{x_1, x_2\}$, and we have a BPA m on X such that $m(X) = 1$, then the value of E_{Zhou} is lower than the value of the Hartley measure, as shown in Section 3. We have also demonstrated that, in the rest of the cases where there is only one focal set, the value of E_{Zhou} is strictly greater than the one obtained with the Hartley measure. It might be an inconsistent behavior.
- Regarding the conflict parts of E_{Zhou} and E_{Cui}, they can be positive in cases where all the focal sets share an element. It is not logical since the conflict in DST corresponds to cases where the information is focused on sets whose intersection is empty. It also occurs with the original Deng entropy [16].
- Finally, the generalization of the modified Deng entropies considered in this work to more general theories than DST is still an open question. As explained in [11], it must be possible to extend a TU measure in DST to more general theories. It is consistent with the *principle of uncertainty invariance* [1]. According to it, "when a representation of uncertainty in one mathematical theory is transformed into its counterpart in another theory, the amount of information must be preserved".

5. Conclusions and future work

In this work, we have analyzed the mathematical properties for uncertainty measures in DST of two modifications of the Deng entropy. The first one of them, E_{Zhou}, takes into consideration the number of alternatives to a higher degree than the original Deng entropy. The second modification, E_{Cui}, also improves the original Deng entropy because it considers the intersections between the focal elements.

This analysis has shown that, as the original Deng entropy, the modified ones only satisfy one of the six essential mathematical properties. Some of these properties are debatable, but other ones are crucial, such as subadditivity, additivity, and monotonicity. Remark that both E_{Zhou} and E_{Cui} are sometimes inconsistent when it is produced an increase or decrease of information, which is very incoherent.

Moreover, it has also shown that the modified Deng entropies considered here also present some questionable behaviors. For example, in both E_{Zhou} and E_{Cui}, the conflict part is positive in situations where all the focal sets share an element. Also, the generalization of these modified Deng entropies to more general theories than DST is not trivial.

For all the above reasons, the modified Deng entropies E_{Zhou} and E_{Cui} do not solve most of the problems of the original one and, consequently, they should be cautiously employed in practical applications.

As future research, the set of required properties that uncertainty measures in DST must satisfy could be revised. As we have said, some properties are debatable. For instance, since the non-specificity is not present in PT, unlike the conflict, it might be reasonable that the range of an uncertainty measure in DST is enlarged.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Klir GJ. Uncertainty and information: foundations of generalized information theory. John Wiley and Sons, Inc; 2005. ISBN 9780471755579. doi:101002/0471755575.

[2] Dempster AP. Upper and lower probabilities induced by a multivalued mapping. Ann Math Stat 1967;38(2):325–39. doi:10.1214/aoms/1177698950.
[3] Shafer G. A mathematical theory of evidence. Princeton university press Princeton; 1976.
[4] Beynon M, Curry B, Morgan P. The Dempster Shafer theory of evidence: an alternative approach to multicriteria decision modelling. Omega 2000;28(1):37–50. doi:10.1016/S0305-0483(99)00033-X.
[5] Denœux T. A k-nearest neighbor classification rule based on Dempster-Shafer theory. Springer Berlin Heidelberg; 2008. p. 737–60. ISBN 978-3-540-44792-4.
[6] Ip HHS, Ng JMC. Human face recognition using Dempster-Shafer theory, vol. 2; 1994. p. 292–5.
[7] Buede DM, Girardi P. A target identification comparison of bayesian and Dempster-Shafer multisensor fusion. IEEE Trans Syst Man Cybern Part A 1997;27(5):569–77. doi:10.1109/3468.618256.
[8] Shannon CE. A mathematical theory of communication. Bell Syst Tech J 1948;27(3):379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
[9] Yager RR. Entropy and specificity in a mathematical theory of evidence. Int J Gen Syst 1983;9(4):249–60. doi:10.1080/03081078308960825.
[10] Klir G, Wierman M. Uncertainty-Based information: elements of generalized information theory. Physica-Verlag HD; 1999.
[11] Abellán J, Masegosa A. Requirements for total uncertainty measures in Dempster-Shafer theory of evidence. Int J Gen Syst 2008;37(6):733–47. doi:10.1080/03081070802082486.
[12] Harmanec D, Klir GJ. Measuring total uncertainty in Dempster-Shafer theory: a novel approach. Int J Gen Syst 1994;22(4):405–19. doi:10.1080/03081079408935225.
[13] Maeda Y, Nguyen HT, Ichihashi H. Maximum entropy algorithms for uncertainty measures. Int J Uncertain Fuzziness Knowl Based Syst 1993;01(01):69–93. doi:10.1142/S021848859300005X.
[14] Huynh V-N, Nakamori Y. Notes on reducing algorithm complexity for computing an aggregate uncertainty measure. IEEE Trans Syst Man CybernPart A 2010;40(1):205–9. doi:10.1109/TSMCA.2009.2030962.
[15] Deng Y. Deng entropy. Chaos Solitons Fractals 2016;91:549–53 110.1016/j.chaos.2016.07.014.
[16] Abellán J. Analyzing properties of Deng entropy in the theory of evidence. Chaos Solitons Fractals 2017;95:195–9. doi:10.1016/j.chaos.2016.12.024.
[17] Zhou D, Tang Y, Jiang W. A modified belief entropy in Dempster-Shafer framework. PLoS ONE 2017;12(5):1–17. doi:10.1371/journal.pone.0176832.
[18] Cui H, Liu Q, Zhang J, Kang B. An improved Deng entropy and its application in pattern recognition. IEEE Access 2019;7:18284–92.
[19] Hartley RVL. Transmission of information1. Bell Syst Tech J 1928;7(3):535–63. doi:10.1002/j.1538-7305.1928.tb01236.x.
[20] Dubois D, Prade H. A note on measures of specificity for fuzzy sets. Int J Gen Syst 1985;10(4):279–83. doi:10.1080/03081078508934893.
[21] Abellán J, Moral S. A non-specificity measure for convex sets of probability distributions. Int J Uncertain Fuzziness Knowl Based Syst 2000;08(03):357–67. doi:10.1142/S0218488500000253.
[22] Abellán J, Klir G, Moral S. Disaggregated total uncertainty measure for credal sets. Int J Gen Syst 2006;35(1):29–44. doi:10.1080/03081070500473490.
[23] Klir GJ, Ramer A. Uncertainty in the Dempster-Shafer theory: a critical re-examination. Int J Gen Syst 1990;18(2):155–66. doi:10.1080/03081079008935135.
[24] Lamata MT, Moral S. Measures of entropy in the theory of evidence. Int J Gen Syst 1988;14(4):297–305. doi:10.1080/03081078808935019.
[25] Maeda Y, Ichihashi H. An uncertainty measure with monotonicity under the random set inclusion. Int J Gen Syst 1993;21(4):379–92. doi:10.1080/03081079308945088.
[26] Abellán J, Moral S. Upper entropy of credal sets. applications to credal classification. International Journal of Approximate Reasoning 2005;39(2–3):235–55. doi:10.1016/j.ijar.2004.10.001. Imprecise probabilities and their applications
[27] Kang B, Deng Y. The maximum Deng entropy. IEEE Access 2019;7:120758–65. doi:10.1109/ACCESS.2019.2937679.
[28] Zhu R, Chen J, Kang B. Power law and dimension of the maximum value for belief distribution with the maximum Deng entropy. IEEE Access 2020;8:47713–19. doi:10.1109/ACCESS.2020.2979060.