Foundations for a new theory
of plausible and paradoxical reasoning

Jean Dezert

Onera, DTIM/IED,
29 Avenue de la Division Leclerc, 92320 Châtillon, France
`Jean.Dezert@onera.fr`

Abstract. This paper brings foundations for a new theory of plausible and paradoxical reasoning and describes a rule of combination of sources of information in a very general framework where information can be both uncertain and paradoxical. In this new theory, the rule of combination which takes into account explicitly both conjunctions and disjunctions of assertions in the fusion process, appears to be more simple and general than the Dempster's rule of combination. Through several simple examples, we show the strong ability of this new theory to solve practical but difficult problems where the Dempster-Shafer theory usually fails.

1 Introduction

The processing of uncertain information has always been a hot topic of research since mainly the 18th century. Up to middle of the 20th century, most theoretical advances have been devoted to the theory of probabilities through the works of eminent mathematicians like J. Bernoulli (1713), A. De Moivre (1718), T. Bayes (1763), P. Laplace (1774), K. Gauss (1823), S. Poisson (1837), E. Borel (1909), R. Fisher (1930), A. Kolmogorov (1933), B. De Finetti (1958), L. Savage (1967), T. Fine (1973), E. Jaynes (1995) to name just few of them. With the development of computer science, the last half of the 20th century has became very prolific for the development of new original theories dealing with uncertainty and imprecise information. Mainly three major theories are available now as alternative of the theory of probabilities for the automatic plausible reasoning in expert systems: the fuzzy set theory developed by L. Zadeh in sixties (1965), the Shafer's theory of evidence in seventies (1976) and the theory of possibilities by D. Dubois and H. Prade in eighties (1985) and, very recently, the unifying *avant-gardiste* neutrosophy theory proposed by F. Smarandache (2000) in [54, 55, 2]. This paper is focused on the development of a new theory of plausible and paradoxical reasoning which can be interpreted as a generalization of the theory of evidence. After a brief presentation of the Dempster-Shafer theory in section 2, we set up the foundations of our new theory in section 3 and discuss about the justification of the new rule of combination of uncertain and paradoxical sources of evidences. Several illustrative examples of the power and the usefulness of our new theory are also presented and compared with the classical results drawn from the classical Dempster-Shafer theory.

2   The Dempster-Shafer theory of evidence

We present here briefly the basis of the Dempster-Shafer theory (DST) or the Mathematical Theory of Evidence (MTE) [47, 10] called also sometimes the theory of probable or evidential reasoning. The DST is usually considered as a generalization of the bayesian theory of subjective probability [52] and offers a simple and direct representation of ignorance. The DST has shown its compatibility with the classical probability theory, with boolean logic and has a feasible computational complexity [44] for problems of small dimension. The DST is a powerful theoretical tool which can be applied for the representation of incomplete knowledge, belief updating, and for combination of evidence [41, 19] through the Demspter-Shafer's rule of combination presented in the following. The Dempster-Shafer model of representation and processing of uncertainty has led to a huge number of practical applications in a wide range of domains (for example for the pattern classification [12], the integration of knowledge from heterogeneous sources for object identification and tracking [46], autonomous navigation [14], technical and medical diagnosis under unreliable measuring devices, information retrieval, multisensor image segmentation, network reliability computation, safety control in large plants, map construction and maintenance, just to mention a few).

2.1   Basic belief masses

Let $\Theta = \{\theta_i, i = 1, \ldots, n\}$ be a finite discrete set of *exhaustive* and *exclusive* elements (hypotheses) called elementary elements. Θ has been called the frame of discernment of hypotheses or universe of discourse by G. Shafer. The cardinality (number of elementary elements) of Θ is denoted $|\Theta|$. The power set $\mathcal{P}(\Theta)$ of Θ which is the set of all subsets of Θ is usually denoted by $\mathcal{P}(\Theta) = 2^\Theta$ because its cardinality is exactly $2^{|\Theta|}$. Any element of 2^Θ is then a composite event (disjunction) of the frame of discernment.

Definition 1. *The DST starts by defining a map associated to a body of evidence* $\mathcal{B}$ *(source of information), called basic belief assignment (bba)*[1] *or information granule* $m(.) : 2^\Theta \to [0, 1]$ *such that*

$$m(\emptyset) = 0 \tag{1}$$

$$\sum_{A \in 2^\Theta} m(A) \equiv \sum_{A \subseteq \Theta} m(A) = 1 \tag{2}$$

$m(.)$ represents the strength of some evidence provided by the source of information under consideration. Condition (1) reflects the fact that no belief ought to be committed to $\emptyset$ and condition (2) reflects the convention that one's total belief has measure one [47]. $m(A)$ corresponds to the measure of the partial

---

[1] This terminology suggested by Professor Philippe Smets to the author appears to be less confusing than the basic probability assignment terminology (bpa) originally adopted by Glenn Shafer.

belief that is committed *exactly* to A (degree of truth supported exactly by A) by the body of evidence $\mathcal{B}$ but not the total belief committed to A. All subsets A for which $m(A) > 0$ are called focal elements of m. The set of all focal elements of $m(.)$ is called the core $\mathcal{K}(m)$ of m. Note that $m(A_1)$ and $m(A_2)$ can be both equal to zero even if $m(A_1 \cup A_2) \neq 0$. Even more peculiar, note that $A \subset B \nRightarrow m(A) < m(B)$ (i.e. $m(.)$ is not monotone to inclusion). Hence, the bba $m(.)$ is in general different from a probability distribution $p(.)$.

Example 1. Consider $\Theta = \{\theta_1, \theta_2, \theta_3\}$, then $2^{\Theta} = \{\emptyset, \theta_1, \theta_2, \theta_3, \theta_1 \cup \theta_2, \theta_1 \cup \theta_3, \theta_2 \cup \theta_3, \theta_1 \cup \theta_2 \cup \theta_3\}$. An information granule $m(.)$ on this frame of discernment Θ could be defined as

$$
\begin{aligned}
m(\emptyset) &\triangleq 0 & m(\theta_1 \cup \theta_2 \cup \theta_3) &= 0.05 \\
m(\theta_1) &= 0.40 & m(\theta_1 \cup \theta_2) &= 0.10 \\
m(\theta_2) &= 0.20 & m(\theta_2 \cup \theta_3) &= 0.10 \\
m(\theta_3) &= 0.05 & m(\theta_1 \cup \theta_3) &= 0.10
\end{aligned}
$$

In this particular example $\mathcal{K}(m) = \{\theta_1, \theta_2, \theta_3, \theta_1 \cup \theta_2, \theta_1 \cup \theta_3, \theta_2 \cup \theta_3, \theta_1 \cup \theta_2 \cup \theta_3\}$ and note that $\theta_1 \subset \{\theta_1 \cup \theta_2\}$ with $m(\theta_1) > m(\theta_1 \cup \theta_2)$.

2.2 Belief functions

Definition 2. *To measure the total belief committed to $A \in 2^{\Theta}$, Glenn Shafer has defined the belief (or credibility) function $Bel(.) : 2^{\Theta} \rightarrow [0, 1]$ associated with bba $m(.)$ as*

$$
Bel(A) = \sum_{B \subseteq A} m(B) \tag{3}
$$

$Bel(A)$ summarizes all our reasons to believe in A (i.e. the lower probability to believe in A). More generally, a belief function $Bel(.)$ can be characterized without reference to the information granule $m(.)$ if $Bel(.)$ satisfies the following three conditions $\forall n > 0, \forall A_1, \ldots, A_n \subset \Theta$,

$$
Bel(\Theta) = 1 \tag{4}
$$

$$
Bel(\emptyset) = 0 \tag{5}
$$

$$
Bel(A_1 \cup \ldots \cup A_n) \geq \sum_{\substack{I \subset \{1, \ldots, n\} \\ I \neq \emptyset}} (-1)^{|I|+1} Bel(\bigcap_{i \in I} A_i) \tag{6}
$$

For any given belief function $Bel(.)$, one can always associate an unique information granule $m(.)$, called the Möbius inverse of the belief function [42], and defined by [47]

$$
\forall A \subseteq \Theta, \qquad m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} Bel(B) \tag{7}
$$

Definition 3. *The vacuous belief function having $Bel(\Theta) = 1$ but $Bel(A) = 0$ for all $A \neq \Theta$ describes the full ignorance on the frame of discernment Θ. The corresponding bba $m_v(.)$ is such that $m_v(\Theta) = 1$ and $m_v(A) = 0$ for all $A \neq \Theta$.*

Proposition 1. *For any given belief function $Bel(.)$ defined on Θ, one has*

$$\forall A, B \subseteq \Theta, \ max(0, Bel(A) + Bel(B) - 1) \leq Bel(A \cap B) \leq min(Bel(A), Bel(B))$$

Definition 4. *Any belief function satisfying $Bel(\emptyset) = 0$, $Bel(\Theta) = 1$ and $Bel(A \cup B) = Bel(A) + Bel(B)$ whenever $A, B \subset \Theta$ and $A \cap B = \emptyset$ is called a bayesian belief function.*

In this case, (6) coincides exactly with the well-known Poincaré's equality

$$P\{A_1 \cup \ldots \cup A_n\} = \sum_{\substack{I \subset \{1,\ldots,n\} \\ I \neq \emptyset}} (-1)^{|I|+1} P\{\bigcap_{i \in I} A_i\} \tag{8}$$

A probability function $P(.)$ can then be interpreted as a particular Dempster-Shafer's belief function. In this sense, the Dempster-Shafer theory is usually considered as a generalization of the probability theory.

Proposition 2. *If $Bel(.)$ is a bayesian belief function, then all focal elements are only single points of $\mathcal{P}(\Theta)$. The basic belief assignement $m(.)$ commits a positive number $m(\theta_i)$ only to some elementary $\theta_i \in \Theta$ (possibly all θ_i) and zero to all possible disjunctions of $\theta_1, \ldots, \theta_n$. In other words, there exists a bayesian bba $m(.) : \Theta \rightarrow [0, 1]$ such that*

$$\sum_{\theta_i \in \Theta} m(\theta_i) = 1 \qquad and \qquad \forall A \subseteq \Theta, \quad Bel(A) = \sum_{\theta_i \in A} m(\theta_i) \tag{9}$$

2.3   Plausibility functions

Since the degree of belief $Bel(A)$ does not reveal to what extent one believes its negation A^c, G. Shafer has introduced the degree of doubt of A as the total belief of A^c. The degree of doubt is less useful than the plausibility $Pl(A)$ of A which measures the total belief mass that can move into A (interpreted sometimes as the *upper probability* of A).

Definition 5. *More precisely, the plausibility $Pl(A)$ of any assertion $A \subset 2^\Theta$ is defined by*

$$Pl(A) \triangleq 1 - Bel(A^c) = \sum_{B \subseteq \Theta} m(B) - \sum_{B \subseteq A^c} m(B) = \sum_{B \cap A \neq \emptyset} m(B) \tag{10}$$

The dual of (6) implies $\forall n > 0, \forall A_1, \ldots, A_n \subset \Theta$,

$$Pl(A_1 \cap \ldots \cap A_n) \leq \sum_{\substack{I \subset \{1,\ldots,n\} \\ I \neq \emptyset}} (-1)^{|I|+1} Pl(\bigcup_{i \in I} A_i) \tag{11}$$

The comparison of (3) with (10) indicates that $\forall A \subseteq \Theta$, $Bel(A) \leq Pl(A)$.

Proposition 3. *For any given plausibility function Pl(.) defined on frame of discernment Θ, the following inequality holds [47]*

$$\forall A, B \subseteq \Theta, \ max(Pl(A), Pl(B)) \leq Pl(A \cup B) \leq min(1, Pl(A) + Pl(B)) \quad (12)$$

Let Θ be a given frame of discernment and $m(.)$ a general bba (neither a vacuous bba, nor a bayesian bba) provided by a body of evidence, then it is always possible to build the following *pignistic*[2] probability [56, 62] (bayesian belief function) by choosing $\forall \theta_i \in \Theta, P\{\theta_i\} = \sum_{B \subseteq \Theta | \theta_i \in B} \frac{1}{|B|} m(B)$. In such case, one always has

$$\forall A \subseteq \Theta, \qquad \mathrm{Bel}(A) \quad \leq \quad [P(A) = \sum_{\theta_i \in A} P\{\theta_i\}] \quad \leq \quad \mathrm{Pl}(A) \qquad (13)$$

Since $\mathrm{Bel}(A)$ summarizes all our reasons to believe in A and $\mathrm{Pl}(A)$ expresses how much we should believe in A if all currently unknown were to support A, the true belief in A is somewhere in the interval $[\mathrm{Bel}(A), \mathrm{Pl}(A)]$. Now suppose that the true value of a parameter under consideration is known with some uncertainty $[\mathrm{Bel}(A), \mathrm{Pl}(A)] \subseteq [0, 1]$, then its corresponding bba $m(A)$ can always be constructed by choosing

$$m(A) = \mathrm{Bel}(A) \qquad m(A \cup A^c) = \mathrm{Pl}(A) - \mathrm{Bel}(A) \qquad m(A^c) = 1 - \mathrm{Pl}(A)$$

2.4   The Dempster's rule of combination

Glenn Shafer has proposed the Dempster's rule of combination (orthogonal summation), symbolized by the operator $\oplus$, to combine two so-called distinct bodies of evidences $\mathcal{B}_1$ and $\mathcal{B}_2$ over the same frame of discernment Θ. Let $\mathrm{Bel}_1(.)$ and $\mathrm{Bel}_2(.)$ be two belief functions over the same frame of discernment Θ and $m_1(.)$ and $m_2(.)$ their corresponding bba masses. The combined global belief function $\mathrm{Bel}(.) = \mathrm{Bel}_1(.) \oplus \mathrm{Bel}_2(.)$ is obtained from the combination of the information granules $m_1(.)$ and $m_2(.)$ as follows: $m(\emptyset) = 0$ and for any $C \neq \emptyset$ and $C \subseteq \Theta$,

$$m(C) \triangleq [m_1 \oplus m_2](C) = \frac{\sum_{A \cap B = C} m_1(A) m_2(B)}{\sum_{A \cap B \neq \emptyset} m_1(A) m_2(B)}$$

$$= \frac{\sum_{A \cap B = C} m_1(A) m_2(B)}{1 - \sum_{A \cap B = \emptyset} m_1(A) m_2(B)} \qquad (14)$$

$\sum_{A \cap B = C}$ represents the sum over all $A, B \subseteq \Theta$ such that $A \cap B = C$ (the interpretation for other summation notations follows directly by analogy). The orthogonal sum $m(.)$ is a proper bba if $K \triangleq 1 - k = 1 - \sum_{A \cap B = \emptyset} m_1(A) m_2(B) \neq 0$. If $K = 0$, which means $\sum_{A \cap B = \emptyset} m_1(A) m_2(B) = 1$ then orthogonal sum $m(.)$ *does not exist* and the bodies of evidences $\mathcal{B}_1$ and $\mathcal{B}_2$ are said to be totally (flatly) contradictory or in *full contradiction*. Such case arises whenever the cores of $\mathrm{Bel}_1(.)$

---

[2] We adopt here the historical definition of the pignistic probability coined by P. Smets (but already proposed independently in the meantime in [13]). New pignistic probabilities have recently been proposed by J. Sudano in [65, 66].

and $\mathrm{Bel}_2(.)$ are disjoint or equivalently when there exists $A \subset \Theta$ such that $\mathrm{Bel}_1(A) = 1$ and $\mathrm{Bel}_2(A^c) = 1$. The same problem of existence has already been pointed out in the presentation of the optimal bayesian fusion rule in [15]. The quantity $\log 1/K$ is called the *weight of conflict* between the bodies of evidences $\mathcal{B}_1$ and $\mathcal{B}_2$. It is easy to show that the Dempster's rule of combination is commutative ($m_1 \oplus m_2 = m_2 \oplus m_1$) and associative ($[m_1 \oplus m_2] \oplus m_3 = m_1 \oplus [m_2 \oplus m_3]$). The vacuous belief function such that $m_v(\Theta) = 1$ and $m_v(A) = 0$ for $A \neq \Theta$ is the identity element for $\oplus$ fusion operator, i.e. $m_v \oplus m = m \oplus m_v \equiv m$. If $\mathrm{Bel}_1(.)$ and $\mathrm{Bel}_2(.)$ are two combinable belief functions and if $\mathrm{Bel}_1(.)$ is bayesian, then $\mathrm{Bel}_1 \oplus \mathrm{Bel}_2$ is a bayesian belief function.

This *ad hoc* rule of combination proposed by G. Shafer in [47] (see also discussion in [49]) has been strongly criticized in the past decades but is now accepted since the axiomatic of the transferable belief model developed by Smets in [57, 18, 27, 60, 61] from an idea initiated by Cheng and Kashyap in [5]. Another approach for the justification of Dempster's rule of combination based on the Mathematical Theory of Hint (MTH) has been also proposed by Kohlas in [32]. Discussions on justifications and interpretations of the DST and the Dempster's rule of combination can be also found in [17, 30, 31, 33, 41, 43, 69]. An interesting discussion on the justification of Dempster's rule of combination from the information entropy viewpoint based on the measurement projection and balance principles can be found in [67]. Connection of the DST with the fuzzy set theory can be found in [3, 63] and relationship between foundations of the fuzzy set theory and the probability theory is discussed in [8]. The relationship between experimental observations and the DST belief functions is currently a hot topic of research. Several models have been developed for fitting belief functions with experimental data. A very recent detailed presentation and discussion on this problem can be found in [70].

In the bayesian framework, if we consider M independent sources of information (bodies of evidence) $\mathcal{B}_1, \ldots, \mathcal{B}_M$ providing M *subjective* probability functions $P_1\{.\}, \ldots, P_M\{.\}$ over the same space Θ, then the optimal bayesian fusion rule is given by (see [15] for a general and a theoretical justification)

$$P_{1,\ldots,M}\{\theta_i\} \triangleq [P_1 \oplus \ldots \oplus P_M]\{\theta_i\} = \frac{p_i^{1-M} \prod_{m=1,M} P_m\{\theta_i\}}{\sum_{i=1,n} p_i^{1-M} \prod_{m=1,M} P_m\{\theta_i\}} \qquad (15)$$

where p_i is the prior probability of θ_i. It is easy to check (when the fusion rule is numerically well defined) that this optimal rule of combinations reduces to

$$P_{1,\ldots,M}\{\theta_i\} = [P_1 \oplus \ldots \oplus P_M]\{\theta_i\} = \frac{\prod_{m=1,M} P_m\{\theta_i\}}{\sum_{i=1,n} \prod_{m=1,M} P_m\{\theta_i\}} \qquad (16)$$

if we admit the principle of indifference (by setting all $p_i = 1/n$).

In the last case, one can see a strong similarity between the Dempster's rule and the optimal bayesian fusion rule. Actually, the classical bayesian inference $P\{A \mid B\} = P\{B \mid A\}P\{A\}/P\{B\}$ can be interpreted as a special case of bayesian rule of combination (16) between two sources of information (between prior and posterior information). The Dempster's and Bayes' fusion rules coincide exactly when $m_1(.)$ and $m_2(.)$ become bayesian basic probability mass assignments and if we accept the principle of indifference within the optimal bayesian fusion rule.

The complexity of DS rule of combination is important in general (when we deal with large frames of discernment) since the computational burden for finding all pairs A and B of subsets of Θ such that $A \cap B = C$ is $o(2^{|\Theta|-|C|} \times 2^{|\Theta|-|C|})$ becomes a huge number. For example, if $|\Theta| = 10$ and $|C| = 2$, we will have $o(2^{16}) = o(65536)$ tests to do to find $\{A \cap B | A \cap B = C\}$. But there exists fortunately a fast Möbius transform which allows an efficient implementation of DS rule of combination [25, 26] to deal with problems of high dimension.

Example 2. A simple example of the Dempster's rule of combination

Consider the simple frame of discernment $\Theta = \{S(\text{unny}), R(\text{ainy})\}$ about the true nature of the weather at a given location L for the next day and let consider two independent bodies of evidence $\mathcal{B}_1$ and $\mathcal{B}_2$ providing the following weather forecasts at L

$$m_1(S) = 0.80 \qquad m_1(R) = 0.12 \qquad m_1(S \cup R) = 0.08$$

$$m_2(S) = 0.90 \qquad m_2(R) = 0.02 \qquad m_2(S \cup R) = 0.08$$

The Dempster's rule yields the following result (where $K = 1 - 0.108 - 0.016$)

$$m(S) = (m_1 \oplus m_2)(S) = (0.72 + 0.072 + 0.064)/K \approx 0.977$$
$$m(R) = (m_1 \oplus m_2)(R) = (0.0024 + 0.0096 + 0.0016)/K \approx 0.016$$
$$m(S \cup R) = (m_1 \oplus m_2)(S \cup R) = 0.0064/K \approx 0.007$$

Hence, in this example, the fusion of the two sources of evidence reinforces the belief that tomorrow will be a sunny day at location L (assuming that both bodies of evidence are equally reliable).

Example 3. Another simple but disturbing example

In 1982, Lofti Zadeh [74] has given to Philippe Smets during a dinner at Acapulco, the following example of a use of the Dempster's rule which shows an unexpected result drawn from the DST. Two doctors examine a patient and agree that it suffers from either meningitis (M), concussion (C) or brain tumor (T). Thus $\Theta = \{M, C, T\}$. Assume that the doctors agree in their low expectation of a tumor, but disagree in likely cause and provide the following diagnosis

$$m_1(M) = 0.99 \quad m_1(T) = 0.01 \quad \text{and} \quad m_2(C) = 0.99 \quad m_2(T) = 0.01$$

If we now combine belief functions using Dempster's rule of combination, one gets the unexpected final conclusion $m(T) = \frac{0.0001}{1-0.0099-0.0099-0.9801} = 1$ which means that the patient suffers with certainty from brain tumor !!!. This unexpected result arises from the fact that the two bodies of evidence (doctors) agree that patient does not suffer from tumor but are in almost full contradiction for the other causes of the disease. This very simple but practical example shows the limitations of practical use of the DST for automated reasoning. Some extreme caution on the degree of conflict of the sources must always be taken before taking a final decision based on the Dempster's rule of combination. A justification of non effectiveness of the Dempster's rule in such kind of example based on an information entropy argument has already been reported in [67].

Example 4. Blackman's example

Let's consider now the example[3] provided by Samuel Blackman in [4] (pp. 207-209). Consider only two attribute types corresponding to the frame of discernment $\Theta = \{\theta_1, \theta_2\}$ and the assignment problem for a single observation and two tracks (T_1 and T_2). Assume now the following two predicted bba for the two tracks:

$$m_{T_1}(\theta_1) = 0.5 \qquad m_{T_1}(\theta_2) = 0.5 \qquad m_{T_1}(\theta_1 \cup \theta_2) = 0$$

$$m_{T_2}(\theta_1) = 0.1 \qquad m_{T_2}(\theta_2) = 0.1 \qquad m_{T_2}(\theta_1 \cup \theta_2) = 0.8$$

Now assume to receive the new following bba drawn from attribute observation Z of the system

$$m_Z(\theta_1) = 0.5 \qquad m_Z(\theta_2) = 0.5 \qquad m_Z(\theta_1 \cup \theta_2) = 0$$

The observation bba $m_Z(.)$ fits perfectly with the predicted bba $m_{T_1}(.)$ whereas $m_Z(.)$ has some disagreement with the predicted bba $m_{T_2}(.)$. If we use the DST to solve this very simple assignment problem between the observation and several predicted bba, one gets from the DS rule of combination exactly the same result, i.e. for $m_{T_1 Z} \triangleq m_{T_1} \oplus m_Z$ and $m_{T_2 Z} \triangleq m_{T_2} \oplus m_Z$:

$$m_{T_1 Z}(\theta_1) = 0.5 \qquad m_{T_1 Z}(\theta_2) = 0.5 \qquad m_{T_1 Z}(\theta_1 \cup \theta_2) = 0$$

$$m_{T_2 Z}(\theta_1) = 0.5 \qquad m_{T_2 Z}(\theta_2) = 0.5 \qquad m_{T_2 Z}(\theta_1 \cup \theta_2) = 0$$

From these two same results only, it is impossible to find the correct solution of this simple assignment problem. Moreover the weights of conflict between sources for the two combinations of evidences are respectively equal to

$$k_{T_1 Z} = 0.5 \qquad \text{for the fusion} \quad m_{T_1} \oplus m_Z$$

$$k_{T_2 Z} = 0.1 \qquad \text{for the fusion} \quad m_{T_2} \oplus m_Z$$

[3] This example has been pointed out to the author by Dr. Albena Tchamova from CLPP, Bulgarian Academy of Sciences during NM&A 02 Conference in Borovetz, Bulgaria, Aug 20-24, 2002

Therefore the resultant conflict terms provides a larger discrepancy between observation bba m_Z with the predicted bba m_{T_1} than with the predicted bba m_{T_2}, despite their bba are equal. Within such conditions, the search for the minimum weight of conflict between sources cannot be taken as a reliable solution for the assignment problem. To solve this anomaly, S. Blackman has proposed to use a relative, rather than an absolute, attribute likelihood function as follows

$$L(Z \mid T) \triangleq (1 - k_{TZ})/(1 - k_{TZ}^{\min})$$

where $k_{TZ}^{\min}$ is the minimum conflict factor that could occur for either the observation Z or the track T in the case of *perfect* assignment (when $m_Z(.)$ and $m_T(.)$ coincide). By adopting this relative likelihhod function, one gets

$$L(Z \mid T_1) = (1-0.5)/(1-0.5) = 1 \quad \text{and} \quad L(Z \mid T_2) = (1-0.1)/(1-0.02) = 0.92$$

Using the Blackman's approach, there is now a larger likelihood associated with the first assignment (hence the right assignment solution can be obtained now based on the max likelihood criteria) but the difference between the two likelihood values is not so big .... As reported by S. Blackman in [4], *more study in this area is required.* Dr. Tchamova has recently proposed, in a private communication to the author, to use the city-block and Euclidean distances $d_1(T, TZ) = \sum_{A \in 2^\Theta} \mid m_T(A) - m_{TZ}(A) \mid$ or $d_2(T, TZ) = \sqrt{\sum_{A \in 2^\Theta} [m_T(A) - m_{TZ}(A)]^2}$ to measure the closeness between $m_{T_1 Z}$ and m_{T_1} and between $m_{T_2 Z}$ and m_{T_2} and then to choose the assignment which corresponds to the minimum distance. Using her approach, one gets

$$d_1(T_1, T_1 Z) = d_2(T_1, T_1 Z) = 0 \qquad d_1(T_2, T_2 Z) = 1.6 \qquad d_2(T_2, T_2 Z) \simeq 0.98$$

The Tchamova's approach can therefore solve the *anomaly* of the DS result in this assignment problem.

Let's consider now the previous predicted gbba $m_{T_1}(.)$ and $m_{T_2}(.)$ but with an observation bba which agrees with $m_{T_2}(.)$ so that $Z \leftrightarrow T_2$ becomes now the correct assignment we are looking for. In other words, let's consider

$$m_Z(\theta_1) = 0.1 \qquad m_Z(\theta_2) = 0.1 \qquad m_Z(\theta_1 \cup \theta_2) = 0.8 \qquad m_Z(\theta_1 \cap \theta_2) = 0$$

Using the DS rule of combination, we get now the following results

$$m_{T_1 Z}(\theta_1) = 0.5 \quad m_{T_1 Z}(\theta_2) = 0.5 \quad m_{T_1 Z}(\theta_1 \cup \theta_2) = 0$$
$$m_{T_2 Z}(\theta_1) = 0.173 \; m_{T_2 Z}(\theta_2) = 0.173 \; m_{T_2 Z}(\theta_1 \cup \theta_2) = 0.654$$

with resulting conflict factors $k_{T_1 Z} = 0.1$ and $k_{T_2 Z} = 0.02$. From these bba $m_{T_1 Z}(.)$, $m_{T_1 Z}(.)$ and conflict factors $k_{T_1 Z}$, $k_{T_2 Z}$ it is clear that the assignment solution is directly given here by the fusion $m_{T_2} \oplus m_Z$ which has the minimum conflict factor. In this second case, we do not need to look for any additional approach to reach the right solution. Nevertheless, it is still interesting to examine the result of the distance approach in this case.

We get then the following distances:

$$d_1(T_1, T_1Z) = d_2(T_1, T_1Z) = 0 \quad d_1(T_2, T_2Z) = 0.292 \quad d_2(T_2, T_2Z) \simeq 0.1788$$

The decision drawn from the minimum distance criteria will yield here the wrong assignment if this approach is chosen.

Therefore, as seen in this simple example, there is no unique and reliable approach to solve the assignment problem based on DST for both cases. In general, we will always have to look for the suitable approach (minimum conflict, Blackman or Tchamova approaches) which allows us to get (hopefully) the correct solution of the problem. The best approach to use can become rather difficult to find by an automatic inference system depending on the complexity of the assignment problem. We will present at the end of this paper how our new theory of plausible and paradoxical reasoning can help to solve this assignment problem. By using only an unique and simple criteria based on our *generalized entropy like measure*, we will be able to provide the correct solution for the two cases of the assignment problem presented in this example.

2.5   Conditional belief functions

Let $m_B(A) = 1$ if $B \subseteq A$ and $m_B(A) = 0$ for if $B \not\subset A$ (the subset B is the only focal element of Bel_B and its basic belief number is one). Then $\text{Bel}_B(.)$ is a belief function that focuses all of the belief on B (note that Bel_B is not in general a bayesian belief function unless $\mid B \mid = 1$).

Definition 6. *Consider now a belief function Bel defined on Θ and a specific belief function Bel_B, then the orthogonal sum denoted as $Bel(. \mid B) = Bel \oplus Bel_B$ is defined for all $A \subset \Theta$ by [47]*

$$Bel(A \mid B) = \frac{Bel(A \cup B^c) - Bel(B^c)}{1 - Bel(B^c)} \tag{17}$$

and

$$Pl(A \mid B) = \frac{Pl(A \cap B)}{Pl(B)} \tag{18}$$

Proposition 4. *If $Bel(.)$ is a bayesian belief function, then*

$$Bel(A \mid B) = \frac{Bel(A \cap B)}{Bel(B)} = Pl(A \mid B) \tag{19}$$

which coincides exactly with the classical conditional probability

$$P\{A \mid B\} = \frac{P\{A \cap B\}}{P\{B\}} \tag{20}$$

3 A new theory for plausible and paradoxical reasoning

3.1 Introduction

As seen in the previous Zadeh's disturbing example, the use of the DST must be done only with extreme caution if one has to take a final and important decision from the result of the Dempster's rule of combination. In most of practical applications based on the DST, some ad-hoc or heuristic recipes must be added to the fusion process to correctly manage or reduce the possibility of high degree of conflict between sources. Otherwise, the fusion results lead to a very dangerous conclusion (or cannot provide a reliable result at all). Even if the DST provides fruithfull results in many applications (mainly in artificial intelligence and systems expert areas) in past decades, we strongly argue that this theory is still too limited because it is based on the two following restrictive constraints :

C1- The DST considers a discrete and finite frame of discernment based on a set of exhaustive and exclusive elementary elements.

C2- The bodies of evidence are assumed independent (each source of information does not take into account the knowledge of other sources) and provide a belief function on the power set 2^{Θ}.

These two constraints are very strong in many practical problems involving uncertain and probable reasoning and dealing with fusion of uncertain,imprecise and paradoxical information. A discussion about this important remark had already been discussed earlier in [34, 35, 45]. In [45], the author has proposed a new partitioning management technique to overcome mainly the C2 constraint. The first constraint is very severe actually since it does not allow paradoxes on elements of the frame of discernment Θ. The DST accepts as foundation the commonly adopted principle of the third exclude. Even if, at the first glance, it makes sense in the traditional classical thought, we develop here a new theory which does not accept this principle of the third exclude and accepts and deals with paradoxes. This is the main purpose and innovation of our new theory referred as the DSmT (standing for Dezert-Smarandache Theory of paradoxical reasoning) in [55].

The constraint C1 assumes that each elementary hypothesis of the frame of discernment Θ is finely and precisely well defined and we are able to discriminate between all elementary hypotheses without ambiguity and difficulty. We argue that this constraint is too limited and that it is not always possible in practice to choose and define a frame of discernment satisfying C1 even for some very simple problems wherein each elementary hypothesis corresponds to a fuzzy or vague concept or attributes. In such cases, the elementary elements of the *frame of discernment* cannot be precisely separated without ambiguity such that no refinement of the frame of discernment satisfying the first constraint is possible.

Example 5. As a simple example, consider an armed robbery situation having a witness and the frame of discernment (associated to the possible size of the thief)

having only two elementary imprecise classes $\Theta = \{\theta_1 = \text{small}, \theta_2 = \text{tall}\}$. An investigator asks the witness about the size of the thief and the witness declares that the thief was tall with bba number $m(\theta_2) = 0.80$, small with bba number $m(\theta_1) = 0.15$ and is uncertain (either tall or small) with $m(\theta_1 \cup \theta_2) = 0.05$. The investigator will have to deal only with this information although the smallness and the tallness have not been precisely defined. The use of this testimony by the investigator (having in other side extra-information about the thief from other sources of information) to infer on the true size of the thief is delicate especially with the important missing information about the size of the witness (who could be either a basketball player, a dwarf or most probably has a size on the average as you and me - assuming you are neither a dwarf nor a basketball player. These both hypotheses are not incompatible actually since some dwarfs really enjoy to play basketball).

Hence, in many situations, the frame of discernment Θ can only be described in terms of imprecise elements which cannot be clearly separated and which cannot be considered as fully disjoint so that the refinement of initial frame into a new one satisfying C1 is like a graal quest and cannot be accomplished. Our last remark about C1 constraint concerns the universal nature of the frame of discernment. It is clear that, in general, the *same* frame of discernment is interpreted differently by the bodies of evidence or experts. Some subjectivity, or at least some fortuitious biases, on the information provided by a source of information is almost unavoidable, otherwise this would assume, as within the DST, that all bodies of evidence have an objective/universal (possibly uncertain) interpretation or measure of the phenomena under consideration. This vision seems to be too restrictive because usually independent bodies of evidence provide their beliefs about some hypotheses only with respect to their own worlds of knowledge and experience. We don't go deeper here in the techniques of refinements and coarsenings of compatible frame of discernments which is a prerequisite to the Dempster's rule of combination (see [47] for details). We just want to emphasize here that the DST cannot be used at all in all cases where C1 cannot be satisfied and we have more generally to accept the idea to deal directly with paradoxical information.

To convince the reader to accept our radical new way of thought, just think about the true nature of a photon? For experts working in particle physics, photons look like particles, for physicists working in electromagnetic field theory, photons are only considered as electromagnetic waves. Both interpretations are true, there is no unicity on the true nature of the photon and actually a photon holds both aspects which appears as a paradoxe for most of human minds. This notion has been accepted in modern physics only with great difficulty and many vigourous discussions about this fondamental question have held at the beginning of the 20th century between all eminent physicists at that time [40].

The constraint C2 hides also a strong difficulty. To apply the Dempter's rule for two independent bodies of evidence $\mathcal{B}_1$ and $\mathcal{B}_2$, it is necessary that both

frames of discernment Θ_1 and Θ_2 (related to each source $\mathcal{B}_1$ and $\mathcal{B}_2$) have to be compatible and to correspond to the same *universal vision* of the possibilities of the answer of the question under consideration. This constraint is itself very difficult to satisfy actually since each source of information has usually only its own (and maybe biaised) interpretation of elements of frame of discernment. The belief provided by each local source of information mainly depends on the own knowledge frame of the source without reference to the (inaccessible) absolute truth of the space of possibilities. Therefore, C2 is, in many cases, also a too strong hypothesis to accept as foundations for a general theory of probable and paradoxical reasoning. A general theory should include the possibility to deal with evidences arising from different sources of information which have no access to absolute interpretation of the elements of the frame of discernment Θ under consideration. This yields to accept paradoxical information as basis for a new general theory of probable reasoning. Actually we will show in the forthcoming examples that the paradoxical information arising from the fusion of several bodies of evidence is very informative and can be used to help us to take legitimous final decision.

In other words, our new theory can be interpreted as a general and direct extension of probability theory and the Dempster-Shafer theory in the following sense. Let $\Theta = \{\theta_1, \theta_2\}$ be the simpliest frame of discernment involving only two elementary hypotheses (with no more additional assumptions on θ_1 and θ_2), then

- the probability theory deals with basic probability assignments $m(.) \in [0,1]$ such that
$$m(\theta_1) + m(\theta_2) = 1$$

- the Dempster-Shafer theory extends the probability theory by dealing with basic belief assignments $m(.) \in [0,1]$ such that
$$m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) = 1$$

- our general theory extends the two previous theories by accepting the possibility for paradoxical information and deals with new basic belief assignments $m(.) \in [0,1]$ such that
$$m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) + m(\theta_1 \cap \theta_2) = 1$$

3.2   Notion of hyper-power set

Let $\Theta = \{\theta_1, \ldots, \theta_n\}$ be a set of n elementary elements considered as exhaustive which cannot be precisely defined and separated so that no refinement of Θ in a new larger set Θ_{ref} of disjoint elementary hypotheses is possible and let's consider the classical set operators $\cup$ (disjunction) and $\cap$ (conjunction). The exhaustive hypothesis about Θ is not a strong constraint since when $\theta_i, i = 1, n$ does not constitute an exhaustive set of elementary possibilities, we can always

add an extra element θ_0 such that $\theta_i, i = 0, n$ describes now an exhaustive set. We will assume therefore, from now on, that Θ characterizes an exhaustive frame of discernment. Θ will be called a *general* frame of discernment in the sequel to emphasize the fact that Θ does not satisfy the Dempster-Shafer C1 constraint.

Definition 7. *The classical power set $\mathcal{P}(\Theta) = 2^\Theta$ has been defined as the set of all proper subsets of Θ when all elements θ_i are disjoint. We extend here this notion and define now the* hyper-power set D^Θ *as the set of all composite possibilities build from Θ with $\cup$ and $\cap$ operators such that $\forall A \in D^\Theta, B \in D^\Theta, (A \cup B) \in D^\Theta$ and $(A \cap B) \in D^\Theta$.*

Obviously, one would always have $D^\Theta \subset 2^{\Theta_{ref}}$ if the refined power set $2^{\Theta_{ref}}$ could be defined and accessible which is not possible in general as already argued.

The cardinality of D^Θ is majored by 2^{2^n} when $\text{Card}(\Theta) = \mid \Theta \mid = n$. The generation of hyper-power set D^Θ corresponds to the famous Dedekind's problem on enumerating the set of monotone Boolean functions (i.e., functions expressible using only AND and OR set operators) [9]. This problem is also related with the Sperner systems [64, 37] based on finite poset (called also as antichains in literature) [6]. The number of antichains on the n-set Θ are equal to the number of monotonic increasing Boolean functions of n variables, and also the number of free distributive lattices with n generators [20, 22, 28, 29, 38, 53]. Determining these numbers is exactly the Dedekind's problem. The choice of letter D in our notation D^Θ to represent the hyper-power set of Θ is in honour of the great mathematician R. Dedekind. The general solution of the Dedekind's problem (for $n > 10$) has not been found yet. We just know that the cardinality numbers of D^Θ follow the integers of the Dedekind's sequence minus one when $\text{Card}(\Theta) = n$ increases.

Example 6.

1. for $\Theta = \{\}$ (empty set), $D^\Theta = \{\emptyset\}$ and $\mid D^\Theta \mid = 1$
2. for $\Theta = \{\theta_1\}$, $D^\Theta = \{\emptyset, \theta_1\}$ and $\mid D^\Theta \mid = 2$
3. for $\Theta = \{\theta_1, \theta_2\}$, $D^\Theta = \{\emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2, \theta_1 \cap \theta_2\}$ and $\mid D^\Theta \mid = 5$
4. for $\Theta = \{\theta_1, \theta_2, \theta_3\}$,

$$D^\Theta = \{\emptyset, \theta_1, \theta_2, \theta_3,$$
$$\theta_1 \cup \theta_2, \theta_1 \cup \theta_3, \theta_2 \cup \theta_3, \theta_1 \cap \theta_2, \theta_1 \cap \theta_3, \theta_2 \cap \theta_3, \theta_1 \cup \theta_2 \cup \theta_3, \theta_1 \cap \theta_2 \cap \theta_3,$$
$$(\theta_1 \cup \theta_2) \cap \theta_3, (\theta_1 \cup \theta_3) \cap \theta_2, (\theta_2 \cup \theta_3) \cap \theta_1, (\theta_1 \cap \theta_2) \cup \theta_3, (\theta_1 \cap \theta_3) \cup \theta_2, (\theta_2 \cap \theta_3) \cup \theta_1,$$
$$(\theta_1 \cup \theta_2) \cap (\theta_1 \cup \theta_3) \cap (\theta_2 \cup \theta_3)\}$$

and $\mid D^\Theta \mid = 19$

It is not difficult, although tedious, to check that $\forall A \in D^\Theta, B \in D^\Theta, (A \cup B) \in D^\Theta$ and $(A \cap B) \in D^\Theta$ (see appendix for the proof).

The extension to larger frame of discernment is possible but requires a higher computational burden. The general and direct analytic computation of $\mid D^{\Theta} \mid$ for a n-set Θ with $n > 10$ is not known and is still under investigations in the mathematical community. Cardinality numbers $\mid D^{\Theta} \mid$ follow the Dedekind's sequence (minus one), $1, 2, 5, 19, 167, 7580, 7828353, \ldots$ when $\text{Card}(\Theta) = n = 0, 1, 2, 3, 4, 5, 6, \ldots$.

3.3   The general basic belief masses $m(.)$

Definition 8. *Let Θ be a general frame of discernment of the problem under consideration. We define a map $m(.) : D^{\Theta} \to [0, 1]$ associated to a given body of evidence $\mathcal{B}$ which can support paradoxical information, as follows*

$$m(\emptyset) = 0 \qquad and \qquad \sum_{A \in D^{\Theta}} m(A) = 1 \tag{21}$$

The quantity $m(A)$ is called A's *general basic belief number* (gbba) or the general basic belief mass for A.

As in the DST, all subsets $A \in D^{\Theta}$ for which $m(A) > 0$ are called focal elements of $m(.)$ and the set of all focal elements of $m(.)$ is also called the core $\mathcal{K}(m)$ of m.

Definition 9. *The belief and plausibility functions are defined in the same way as in the DST, i.e.*

$$Bel(A) = \sum_{B \in D^{\Theta}, B \subseteq A} m(B) \tag{22}$$

$$Pl(A) = \sum_{B \in D^{\Theta}, B \cap A \neq \emptyset} m(B) \tag{23}$$

Note that, we don't define here explicitly the complementary A^c of a proposition A since $m(A^c)$ cannot be precisely evaluated from $\cup$ and $\cap$ operators on D^{Θ} since we include the possibility to deal with a complete paradoxical source of information such that $\forall A \in D^{\Theta}, \forall B \in D^{\Theta}, m(A \cap B) > 0$. These definitions are compatible with the DST definitions when the sources of information become uncertain but rational (they do not support paradoxical information). We still have $\forall A \in D^{\Theta}, \text{Bel}(A) \leq \text{Pl}(A)$.

3.4   Construction of pignistic probabilities from gbba $m(.)$

The construction of a pignistic probability measure from the general basic belief masses $m(.)$ over D^{Θ} with $|\Theta| = n$ is still possible and is given by the general expression of the form

$$\forall i = 1, \ldots, n \qquad P\{\theta_i\} = \sum_{A \in D^{\Theta}} \alpha_{\theta_i}(A) m(A) \tag{24}$$

where $\alpha_{\theta_i}(A) \in [0, 1]$ are weighting coefficients which depend on the inclusion or non-inclusion of θ_i with respect to proposition A. No general analytic expression for $\alpha_{\theta_i}(A)$ has been derived yet even if $\alpha_{\theta_i}(A)$ can be obtained explicitly for simple examples. When general bba $m(.)$ reduces to classical bba (i.e. the DS bba without paradoxe), then $\alpha_{\theta_i}(A) = \frac{1}{|A|}$ if $\theta_i \subseteq A$ and therefore one gets

$$\forall i = 1, \ldots, n \qquad P\{\theta_i\} = \sum_{A \subseteq \Theta | \theta_i \in A} \frac{1}{|A|} m(A) \qquad (25)$$

We present here an example of a pignistic probabilities reconstruction from a general and non degenerated bba $m(.)$ (i.e. $\nexists A \in D^\Theta$ with $A \neq \emptyset$ such that $m(A) = 0$) over D^Θ.

Example 7. If $\Theta = \{\theta_1, \theta_2\}$ then

$$P\{\theta_1\} = m(\theta_1) + \frac{1}{2} m(\theta_1 \cup \theta_2) + \frac{1}{2} m(\theta_1 \cap \theta_2)$$

$$P\{\theta_2\} = m(\theta_2) + \frac{1}{2} m(\theta_1 \cup \theta_2) + \frac{1}{2} m(\theta_1 \cap \theta_2)$$

Example 8. If $\Theta = \{\theta_1, \theta_2, \theta_3\}$ then

$$\begin{aligned}
P\{\theta_1\} = {} & m(\theta_1) + \frac{1}{2} m(\theta_1 \cup \theta_2) + \frac{1}{2} m(\theta_1 \cup \theta_3) + \frac{1}{2} m(\theta_1 \cap \theta_2) + \frac{1}{2} m(\theta_1 \cap \theta_3) \\
& + \frac{1}{3} m(\theta_1 \cup \theta_2 \cup \theta_3) + \frac{1}{3} m(\theta_1 \cap \theta_2 \cap \theta_3) \\
& + \frac{1/2 + 1/3}{3} m((\theta_1 \cup \theta_2) \cap \theta_3) + \frac{1/2 + 1/3}{3} m((\theta_1 \cup \theta_3) \cap \theta_2) \\
& + \frac{1/2 + 1/2 + 1/3}{3} m((\theta_2 \cup \theta_3) \cap \theta_1) \\
& + \frac{1/2 + 1/2 + 1/3}{5} m((\theta_1 \cap \theta_2) \cup \theta_3) \\
& + \frac{1/2 + 1/2 + 1/3}{5} m((\theta_1 \cap \theta_3) \cup \theta_2) \\
& + \frac{1 + 1/2 + 1/2 + 1/3}{5} m((\theta_2 \cap \theta_3) \cup \theta_1) \\
& + \frac{1/2 + 1/2 + 1/3}{4} m((\theta_1 \cup \theta_2) \cap (\theta_1 \cup \theta_3) \cap (\theta_2 \cup \theta_3))
\end{aligned}$$

$$P\{\theta_2\} = m(\theta_2) + \frac{1}{2}m(\theta_1 \cup \theta_2) + \frac{1}{2}m(\theta_2 \cup \theta_3) + \frac{1}{2}m(\theta_1 \cap \theta_2) + \frac{1}{2}m(\theta_2 \cap \theta_3)$$
$$+ \frac{1}{3}m(\theta_1 \cup \theta_2 \cup \theta_3) + \frac{1}{3}m(\theta_1 \cap \theta_2 \cap \theta_3)$$
$$+ \frac{1/2 + 1/3}{3}m((\theta_1 \cup \theta_2) \cap \theta_3)$$
$$+ \frac{1/2 + 1/2 + 1/3}{3}m((\theta_1 \cup \theta_3) \cap \theta_2)$$
$$+ \frac{1/2 + 1/3}{3}m((\theta_2 \cup \theta_3) \cap \theta_1)$$
$$+ \frac{1/2 + 1/2 + 1/3}{5}m((\theta_1 \cap \theta_2) \cup \theta_3)$$
$$+ \frac{1 + 1/2 + 1/2 + 1/3}{5}m((\theta_1 \cap \theta_3) \cup \theta_2)$$
$$+ \frac{1/2 + 1/2 + 1/3}{5}m((\theta_2 \cap \theta_3) \cup \theta_1)$$
$$+ \frac{1/2 + 1/2 + 1/3}{4}m((\theta_1 \cup \theta_2) \cap (\theta_1 \cup \theta_3) \cap (\theta_2 \cup \theta_3))$$

$$P\{\theta_3\} = m(\theta_3) + \frac{1}{2}m(\theta_1 \cup \theta_3) + \frac{1}{2}m(\theta_2 \cup \theta_3) + \frac{1}{2}m(\theta_1 \cap \theta_3) + \frac{1}{2}m(\theta_2 \cap \theta_3)$$
$$+ \frac{1}{3}m(\theta_1 \cup \theta_2 \cup \theta_3) + \frac{1}{3}m(\theta_1 \cap \theta_2 \cap \theta_3)$$
$$+ \frac{1/2 + 1/2 + 1/3}{3}m((\theta_1 \cup \theta_2) \cap \theta_3) + \frac{1/2 + 1/3}{3}m((\theta_1 \cup \theta_3) \cap \theta_2)$$
$$+ \frac{1/2 + 1/3}{3}m((\theta_2 \cup \theta_3) \cap \theta_1)$$
$$+ \frac{1 + 1/2 + 1/2 + 1/3}{5}m((\theta_1 \cap \theta_2) \cup \theta_3)$$
$$+ \frac{1/2 + 1/2 + 1/3}{5}m((\theta_1 \cap \theta_3) \cup \theta_2) + \frac{1/2 + 1/2 + 1/3}{5}m((\theta_2 \cap \theta_3) \cup \theta_1)$$
$$+ \frac{1/2 + 1/2 + 1/3}{4}m((\theta_1 \cup \theta_2) \cap (\theta_1 \cup \theta_3) \cap (\theta_2 \cup \theta_3))$$

The evaluation of weighting coefficients $\alpha_{\theta_i}(A)$ has been obtained from the geometrical interpretation of the relative contribution of the distinct parts of A with the proposition θ_i under consideration. For example, consider $A = (\theta_1 \cap \theta_2) \cup \theta_3$ which corresponds to the area $a_1 \cup a_2 \cup a_3 \cup a_4 \cup a_5$ on the following Venn diagram.
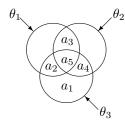
Fig.1 : Representation of $A = (\theta_1 \cap \theta_2) \cup \theta_3 \equiv a_1 \cup a_2 \cup a_3 \cup a_4 \cup a_5$

a_1 which is shared only by θ_3 will contribute to θ_3 with weight 1; a_2 which is shared by θ_1 and θ_3 will contribute to θ_3 with weight 1/2; a_3 which is not shared by θ_3 will contribute to θ_3 with weight 0; a_4 which is shared by θ_2 and θ_3 will contribute to θ_3 with weight 1/2; a_5 which is shared by both θ_1,θ_2 and θ_3 will contribute to θ_3 with weight 1/3. Since moreover, one must have $\forall A \in D^\Theta$ with $m(A) \neq 0$, $\sum_{i=1}^{n} \alpha_{\theta_i}(A)m(A) = m(A)$, it is necessary to normalize $\alpha_{\theta_i}(A)$. Therefore $\alpha_{\theta_1}(A)$, $\alpha_{\theta_2}(A)$ and $\alpha_{\theta_3}(A)$ will be given by

$$\alpha_{\theta_1}(A) = \alpha_{\theta_2}(A) = \frac{1/2 + 1/2 + 1/3}{5} \qquad \alpha_{\theta_3}(A) = \frac{1 + 1/2 + 1/2 + 1/3}{5}$$

All $\alpha_{\theta_i}(A), \forall A \in D^\Theta$ entering in derivation of the *pignistic* probabilities $P\{\theta_i\}$ can be obtained using a similar way.

3.5   General rule of combination of paradoxical sources of evidence

Let's consider now two distinct (but potentially paradoxical) bodies of evidences $\mathcal{B}_1$ and $\mathcal{B}_2$ over the same frame of discernment Θ with belief functions $\text{Bel}_1(.)$ and $\text{Bel}_2(.)$ associated with information granules $m_1(.)$ and $m_2(.)$.

Definition 10. *The combined global belief function $Bel(.) = Bel_1(.) \oplus Bel_2(.)$ is obtained through the combination of the granules $m_1(.)$ and $m_2(.)$ by the simple rule*

$$\forall C \in D^\Theta, \qquad m(C) \triangleq [m_1 \oplus m_2](C) = \sum_{A,B \in D^\Theta, A \cap B = C} m_1(A)m_2(B) \quad (26)$$

Since D^Θ is closed under $\cup$ and $\cap$ operators, this new rule of combination guarantees that $m(.) : D^\Theta \to [0,1]$ is a proper general information granule satisfying (21). The global belief function Bel(.) is then obtained from the granule $m(.)$ through (22). This rule of combination is commutative and associative and can always be used for the fusion of paradoxical or rational sources of information (bodies of evidence). Obviously, the decision process will have to be made with more caution to take the final decision based on the general granule $m(.)$ when internal paradoxical conflicts arise.

It is important to note that any fusion of sources of information generates either uncertainties, paradoxes or *more generally both*. This is intrinsic to the general fusion process itself. For instance, let's consider the frame of discernment $\Theta = \{\theta_1, \theta_2\}$ and the following very simple examples:

Example 9. Consider the rational information granules

$$m_1(\theta_1) = 0.80 \qquad m_1(\theta_2) = 0.20 \qquad m_1(\theta_1 \cup \theta_2) = 0 \qquad m_1(\theta_1 \cap \theta_2) = 0$$

$$m_2(\theta_1) = 0.90 \qquad m_2(\theta_2) = 0.10 \qquad m_2(\theta_1 \cup \theta_2) = 0 \qquad m_2(\theta_1 \cap \theta_2) = 0$$

then

$$m(\theta_1) = 0.72 \qquad m(\theta_2) = 0.02 \qquad m(\theta_1 \cup \theta_2) = 0 \qquad m(\theta_1 \cap \theta_2) = 0.26$$

Example 10. Consider the uncertain information granules

$$m_1(\theta_1) = 0.80 \qquad m_1(\theta_2) = 0.15 \qquad m_1(\theta_1 \cup \theta_2) = 0.05 \qquad m_1(\theta_1 \cap \theta_2) = 0$$

$$m_2(\theta_1) = 0.90 \qquad m_2(\theta_2) = 0.05 \qquad m_2(\theta_1 \cup \theta_2) = 0.05 \qquad m_2(\theta_1 \cap \theta_2) = 0$$

then

$$m(\theta_1) = 0.805 \qquad m(\theta_2) = 0.0175 \qquad m(\theta_1 \cup \theta_2) = 0.0025 \qquad m(\theta_1 \cap \theta_2) = 0.175$$

Example 11. Consider the paradoxical information granules

$$m_1(\theta_1) = 0.80 \qquad m_1(\theta_2) = 0.15 \qquad m_1(\theta_1 \cup \theta_2) = 0 \qquad m_1(\theta_1 \cap \theta_2) = 0.05$$

$$m_2(\theta_1) = 0.90 \qquad m_2(\theta_2) = 0.05 \qquad m_2(\theta_1 \cup \theta_2) = 0 \qquad m_2(\theta_1 \cap \theta_2) = 0.05$$

then

$$m(\theta_1) = 0.72 \qquad m(\theta_2) = 0.0075 \qquad m(\theta_1 \cup \theta_2) = 0 \qquad m(\theta_1 \cap \theta_2) = 0.2725$$

Example 12. Consider the uncertain and paradoxical information granules

$$m_1(\theta_1) = 0.80 \qquad m_1(\theta_2) = 0.10 \qquad m_1(\theta_1 \cup \theta_2) = 0.05 \qquad m_1(\theta_1 \cap \theta_2) = 0.05$$

$$m_2(\theta_1) = 0.90 \qquad m_2(\theta_2) = 0.05 \qquad m_2(\theta_1 \cup \theta_2) = 0.03 \qquad m_2(\theta_1 \cap \theta_2) = 0.02$$

then

$$m(\theta_1) = 0.789 \qquad m(\theta_2) = 0.0105 \qquad m(\theta_1 \cup \theta_2) = 0.0015 \qquad m(\theta_1 \cap \theta_2) = 0.199$$

Note that this general fusion rule can also be used with intuitionist logic in which the sum of bba is allowed to be less than one $(\sum m(A) < 1)$ and with the paraconsistent logic in which the sum of bba is allowed to be greater than one $(\sum m(A) > 1)$ as well. In such cases, the fusion result does not provide in general $\sum m(A) = 1$.

By example, let's consider the fusion of the paraconsistent source $\mathcal{B}_1$ with $m_1(\theta_1) = 0.60$, $m_1(\theta_2) = 0.30$, $m_1(\theta_1 \cup \theta_2) = 0.20$, $m_1(\theta_1 \cap \theta_2) = 0.10$ with the intuitionist source $\mathcal{B}_2$ with $m_2(\theta_1) = 0.50$, $m_2(\theta_2) = 0.20$, $m_2(\theta_1 \cup \theta_2) = 0.10$, $m_2(\theta_1 \cap \theta_2) = 0.10$. In such case, the fusion result of these two sources of information yields the following global paraconsistent bba $m(.)$

$$m(\theta_1) = 0.46 \qquad m(\theta_2) = 0.13 \qquad m(\theta_1 \cup \theta_2) = 0.02 \qquad m(\theta_1 \cap \theta_2) = 0.47$$

which yields $\sum m = 1.08 > 1$.

In practice, for the sake of fair comparison between several alternatives or choices, it is better and simpler to deal with normalized bba to take a final important decision for the problem under consideration. A nice property of the new rule of combination of non-normalized bba is its invariance to the pre- or post-normalization process as we will show right now. In the previous example, the post-normalization of bba $m(.)$ will yield the new bba $m'(.)$

$$m'(\theta_1) = \frac{0.46}{1.08} \approx 0.426 \qquad m'(\theta_2) = \frac{0.13}{1.08} \approx 0.12$$

$$m'(\theta_1 \cup \theta_2) = \frac{0.02}{1.08} \approx 0.019 \qquad m'(\theta_1 \cap \theta_2) = \frac{0.47}{1.08} \approx 0.435$$

The fusion of pre-normalization of bba $m_1(.)$ and $m_2(.)$ will yield the same normalized bba $m'(.)$ since

$$m_1'(\theta_1) = \frac{0.6}{1.2} = 0.50 \qquad m_1'(\theta_2) = \frac{0.3}{1.2} = 0.25$$

$$m_1'(\theta_1 \cup \theta_2) = \frac{0.2}{1.2} \approx 0.17 \qquad m_1'(\theta_1 \cap \theta_2) = \frac{0.1}{1.2} \approx 0.08$$

$$m_2'(\theta_1) = \frac{0.5}{0.9} \approx 0.56 \qquad m_2'(\theta_2) = \frac{0.2}{0.9} \approx 0.22$$

$$m_2'(\theta_1 \cup \theta_2) = \frac{0.1}{0.9} \approx 0.11 \qquad m_2'(\theta_1 \cap \theta_2) = \frac{0.1}{0.9} \approx 0.11$$

$$m'(\theta_1) \approx 0.426 \qquad m'(\theta_2) \approx 0.12 \qquad m'(\theta_1 \cup \theta_2) \approx 0.019 \qquad m'(\theta_1 \cap \theta_2) \approx 0.435$$

It is easy to verify from the general fusion table that the pre or post normalization step yields always the same global normalized bba even for the general case (when $|\Theta| = n$) because the post-normalization constant $\sum m(A)$ is always equal to the product of the two pre-normalization constants $\sum m_1(A)$ and $\sum m_2(A)$.

3.6   Justification of the new rule of combination

Let's consider two bodies of evidence $\mathcal{B}_1$ and $\mathcal{B}_2$ characterized respectively by their bba $m_1(.), m_2(.)$ and their cores $\mathcal{K}_1 = \mathcal{K}(m_1)$, $\mathcal{K}_2 = \mathcal{K}(m_2)$. Following Sun's notation [67], each source of information will be denoted

$$\mathcal{B}_1 = \begin{bmatrix} \mathcal{K}_1 \\ m_1 \end{bmatrix} = \begin{bmatrix} f_1^{(1)} & f_2^{(1)} & \cdots & f_k^{(1)} \\ m_1(f_1^{(1)}) & m_1(f_2^{(1)}) & \cdots & m_1(f_k^{(1)}) \end{bmatrix} \tag{27}$$

$$\mathcal{B}_2 = \begin{bmatrix} \mathcal{K}_2 \\ m_2 \end{bmatrix} = \begin{bmatrix} f_1^{(2)} & f_2^{(2)} & \cdots & f_l^{(2)} \\ m_2(f_1^{(2)}) & m_2(f_2^{(2)}) & \cdots & m_2(f_l^{(2)}) \end{bmatrix} \tag{28}$$

where $f_i^{(1)}, i = 1, k$ are the focal elements of $\mathcal{B}_1$ and $f_j^{(2)}, j = 1, l$ are the focal elements of $\mathcal{B}_2$.

Let's consider now the combined information associated with a new body of evidence $\mathcal{B}$ resulting from the fusion of $\mathcal{B}_1$ and $\mathcal{B}_2$ having bba $m(.)$ with core $\mathcal{K}$. We denote $\mathcal{B}$ as

$$\mathcal{B} \triangleq \mathcal{B}_1 \oplus \mathcal{B}_2 = \begin{bmatrix} \mathcal{K} \\ m \end{bmatrix} = \begin{bmatrix} f_1^{(1)} \cap f_1^{(2)} & f_1^{(1)} \cap f_2^{(2)} & \cdots & f_k^{(1)} \cap f_l^{(2)} \\ m(f_1^{(1)} \cap f_1^{(2)}) & m(f_1^{(1)} \cap f_2^{(2)}) & \cdots & m(f_k^{(1)} \cap f_l^{(2)}) \end{bmatrix} \tag{29}$$

The fusion of the two information granules can be represented with the general table of fusion as follows

$\oplus$	$m_1(f_1^{(1)})$	$m_1(f_2^{(1)})$	...	$m_1(f_i^{(1)})$	...	$m_1(f_k^{(1)})$
$m_2(f_1^{(2)})$	$m(f_1^{(1)} \cap f_1^{(2)})$	$m(f_2^{(1)} \cap f_1^{(2)})$	...	$m(f_i^{(1)} \cap f_1^{(2)})$	...	$m(f_k^{(1)} \cap f_1^{(2)})$
$m_2(f_2^{(2)})$	$m(f_1^{(1)} \cap f_2^{(2)})$	$m(f_2^{(1)} \cap f_2^{(2)})$	...	$m(f_i^{(1)} \cap f_2^{(2)})$	...	$m(f_k^{(1)} \cap f_2^{(2)})$
...	...	...	...	...	...	...
$m_2(f_j^{(2)})$	$m(f_1^{(1)} \cap f_j^{(2)})$	$m(f_2^{(1)} \cap f_j^{(2)})$	...	$m(f_i^{(1)} \cap f_j^{(2)})$	...	$m(f_k^{(1)} \cap f_j^{(2)})$
...	...	...	...	...	...	...
$m_2(f_l^{(2)})$	$m(f_1^{(1)} \cap f_l^{(2)})$	$m(f_2^{(1)} \cap f_l^{(2)})$	...	$m(f_i^{(1)} \cap f_l^{(2)})$	...	$m(f_k^{(1)} \cap f_l^{(2)})$

We look for the optimal rule of combination, i.e. the bba $m(.) = m_1(.) \oplus m_2(.)$ which maximizes the joint entropy of the two information sources. The justification for the Maxent criteria is discussed in [23, 24]. Thus, one has to find $m(.)$ such that [67, 68].

$$\max_m[H(m)] \equiv \max_m \left[-\sum_{i=1}^{k} \sum_{j=1}^{l} m(f_i^{(1)} \cap f_j^{(2)}) \log[m(f_i^{(1)} \cap f_j^{(2)})] \right]$$
$$\equiv -\min_m[-H(m)] \tag{30}$$

satisfying both

1. the measurement projection principle (marginal bba), i.e. $\forall i = 1, \ldots, k$ and $\forall j = 1, \ldots, l$

$$m_1(f_i^{(1)}) = \sum_{j=1}^{l} m(f_i^{(1)} \cap f_j^{(2)}) \text{ and } m_2(f_j^{(2)}) = \sum_{i=1}^{k} m(f_i^{(1)} \cap f_j^{(2)}) \qquad (31)$$

These constraints state that the marginal bba $m_1(.)$ is obtained by the summation over each column of the fusion table and the marginal bba $m_2(.)$ is obtained by the summation over each row of the table of fusion.

2. the measurement balance principle (the sum of all cells of the table of fusion must be unity)

$$\sum_{i=1}^{k} \sum_{j=1}^{l} m(f_i^{(1)} \cap f_j^{(2)}) = 1 \qquad (32)$$

Using the concise notation $m_{ij} \triangleq m(f_i^{(1)} \cap f_j^{(2)})$, the Lagrangian associated with this optimization problem under equality constraints is given by (we consider here the minimization of $-J(m)$ appearing in r.h.s of (30))

$$\mathcal{L}(m, \lambda) = \sum_{i=1}^{k} \sum_{j=1}^{l} m_{ij} \ln[m_{ij}]$$

$$+ \sum_{i=1}^{k} \lambda_i [m_1(f_i^{(1)}) - \sum_{j=1}^{l} m_{ij}] \qquad (33)$$

$$+ \sum_{j=1}^{l} \gamma_j [m_2(f_j^{(2)}) - \sum_{i=1}^{k} m_{ij}] \qquad (34)$$

$$+ \eta [\sum_{i=1}^{k} \sum_{j=1}^{l} m_{ij} - 1] \qquad (35)$$

which can be written more concisely as

$$\mathcal{L}(m, \lambda) = -H(m) + \lambda' \mathbf{g}(m) \qquad (36)$$

where $m = [m_{11} \ m_{12} \ldots \ m_{kl}]'$ and

$$\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_k \\ \gamma_1 \\ \vdots \\ \gamma_l \\ \eta \end{bmatrix} \quad \text{and} \quad \mathbf{g}(m) = \begin{bmatrix} m_1(f_1^{(1)}) - \sum_{j=1}^{l} m_{1j} \\ \vdots \\ m_1(f_k^{(1)}) - \sum_{j=1}^{l} m_{kj} \\ m_2(f_1^{(2)}) - \sum_{i=1}^{k} m_{i1} \\ \vdots \\ m_2(f_l^{(2)}) - \sum_{i=1}^{k} m_{il} \\ \sum_{i=1}^{k} \sum_{j=1}^{l} m_{ij} - 1 \end{bmatrix} \qquad (37)$$

Following the classical method of Lagrange multipliers, one has to find optimal solution (m^*, λ^*) such that

$$\frac{\partial \mathcal{L}}{\partial m}(m^*, \lambda^*) = \mathbf{0} \qquad \text{and} \qquad \frac{\partial \mathcal{L}}{\partial \lambda}(m^*, \lambda^*) = \mathbf{0} \tag{38}$$

The first $k \times l$ equations express the general solution $m[\lambda]$ and the $k + l + 1$ last equations determine λ^* and therefore by substitution into $m[\lambda]$, the optimal solution $m^* = m[\lambda^*]$. One has to solve

$$\frac{\partial \mathcal{L}}{\partial m} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial m_{11}} \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial m_{ij}} \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial m_{kl}} \end{bmatrix} = \begin{bmatrix} \ln(m_{11}) + 1 + \eta - \lambda_1 - \gamma_1 \\ \vdots \\ \ln(m_{ij}) + 1 + \eta - \lambda_i - \gamma_j \\ \vdots \\ \ln(m_{kl}) + 1 + \eta - \lambda_k - \gamma_l \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0} \tag{39}$$

which yields $\forall i, j,$

$$m_{ij} = e^{-\eta - 1} e^{\lambda_i} e^{\gamma_j} \tag{40}$$

and

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial \lambda_1} \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial \lambda_k} \\ \frac{\partial \mathcal{L}}{\partial \gamma_1} \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial \gamma_l} \\ \frac{\partial \mathcal{L}}{\partial \eta} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = \mathbf{0} \Leftrightarrow \begin{bmatrix} e^{-\eta - 1} \sum_{j=1}^{l} e^{\lambda_1} e^{\gamma_j} \\ \vdots \\ e^{-\eta - 1} \sum_{j=1}^{l} e^{\lambda_k} e^{\gamma_j} \\ e^{-\eta - 1} \sum_{i=1}^{k} e^{\gamma_1} e^{\lambda_i} \\ \vdots \\ e^{-\eta - 1} \sum_{i=1}^{k} e^{\gamma_l} e^{\lambda_i} \\ e^{-\eta - 1} \sum_{i=1}^{k} \sum_{j=1}^{l} e^{\gamma_l} e^{\lambda_i} \end{bmatrix} = \begin{bmatrix} m_1(f_1^{(1)}) \\ \vdots \\ m_1(f_k^{(1)}) \\ m_2(f_1^{(2)}) \\ \vdots \\ m_2(f_l^{(2)}) \\ 1 \end{bmatrix} \tag{41}$$

The last constraint in (41) can also be written as

$$e^{-\eta - 1} \sum_{i=1}^{k} \sum_{j=1}^{l} e^{\gamma_l} e^{\lambda_i} = e^{-\eta - 1} \left(\sum_{i=1}^{k} e^{\lambda_i} \right) \left(\sum_{j=1}^{l} e^{\gamma_l} \right) = 1 \tag{42}$$

Now with basic algebraic manipulation, the optimal global bba m_{ij} $\forall i, j$ we are searching for, can be expressed as

$$m_{ij} = e^{-\eta-1}e^{\lambda_i}e^{\gamma_j}$$

$$= e^{-\eta-1}e^{\lambda_i}e^{\gamma_j} \times \overbrace{e^{-\eta-1}(\sum_{i=1}^{k} e^{\lambda_i})(\sum_{j=1}^{l} e^{\gamma_l})}^{1}$$

$$= \underbrace{(e^{-\eta-1}e^{\lambda_i} \sum_{j=1}^{l} e^{\gamma_l})}_{m_1(f_i^{(1)})} \underbrace{(e^{-\eta-1}e^{\gamma_j} \sum_{i=1}^{k} e^{\lambda_i})}_{m_2(f_j^{(2)})}$$

Thus, the solution of the maximisation of the joint entropy is obtained by choosing $\forall i, j$

$$m_{ij} = m(f_i^{(1)} \cap f_j^{(2)}) = m_1(f_i^{(1)})m_2(f_j^{(2)}) \tag{43}$$

Since it may exist several combinations yielding to the same focal element, the bba of all focal elements equal to $f_i^{(1)} \cap f_j^{(2)}$ over the fusion space is

$$m(f_i^{(1)} \cap f_j^{(2)}) = \sum_{i,j} m_1(f_i^{(1)})m_2(f_j^{(2)}) \tag{44}$$

which coincides exactly with the new rule of combination expressed previously.

3.7 The generalized entropy like measure of a source

The evaluation of the entropy $H(m)$ of a given source from the direct extension of its classical definition, with convention (see [7]) $0 \ln(0) = 0$ and with bba $m(.)$, i.e.

$$H(m) = - \sum_{A \in D^\Theta} m(A) \ln(m(A))$$

seems to not be the best measure for the self-information of a general (uncertain and paradoxical) source of information because it does not catch the intrinsic informational strength (i.i.s. for short) $s(A)$ of each proposition A involved in the evaluation of the entropy of the source. An extension of the classical entropy in the DST framework had already been proposed in 1983 by R. Yager based on the weight of conflict between the belief function Bel and the certain support function Bel_A focused on each proposition A (see [71] for details).

Definition 11. *In the classical definition (based only on a probability measure), one always has $s(A) \equiv |A| = 1$. This does not hold in our general theory of plausible and paradoxical reasoning and we propose to generalize the notion of entropy in the following manner to measure the self-information of a general source of information:*

$$H_g(m) = -\sum_{A\in D^\Theta} \frac{1}{s(A)} m(A) \ln(\frac{1}{s(A)} m(A)) \qquad (45)$$

$H_g(m)$ will be called from now on the *generalized entropy* of the source associated with gbba $m(.)$. This definition is coherent with the definition of the classical entropy whenever the gbba $m(.)$ reduces to a basic probability assignment. However in the general case, $H_g(m)$ does not satisfy the properties of the classical entropy (see chap. 1 in [21]). Nevertheless, this generalized entropy like measure can be useful in practice to solve important problems as it will be seen through next examples. This general definition introduces the intrinsic informational strength (called also here the hyper-cardinality) $s(A)$ of a general (irreducible) proposition A which can be derived from the two following important rules

$$s\left(\bigcup_{i=1,n} B_i\right) = s\left(B_1 \cup \ldots \cup B_n\right) = \frac{\sum_{i=1,n} 1/s\left(B_i\right)}{\prod_{i=1,n} 1/s\left(B_i\right)} \qquad (46)$$

$$s\left(\bigcap_{i=1,n} B_i\right) = s\left(B_1 \cap \ldots \cap B_n\right) = \frac{\prod_{i=1,n} s\left(B_i\right)}{\sum_{i=1,n} s\left(B_i\right)} \qquad (47)$$

It is very important to note that these rules apply only on irreducible propositions (logical atoms) A. A proposition A is said to be irreducible (or equivalently has a compact form) if and only if it does not admit other equivalent form with a smaller number of operands and operators. For example $(\theta_1 \cup \theta_3) \cap (\theta_2 \cup \theta_3)$ is not an irreducible proposition since it can be reduced to its equivalent logical atom $(\theta_1 \cap \theta_2) \cup \theta_3$. To compute the i.i.s. $s(A)$ of any proposition A using the rules (46) and (47), the proposition has first to be reduced to its minimal representation (irreducible form).

Example 13. Here are few examples of the value of the hyper-cardinality for some elementary and composite irreducible propositions A. We recall that θ_i involved in A are singletons such that $|\theta_i| = 1$.

$A = \theta_1 \cup \theta_2 \Rightarrow s(A) = 2$

$A = \theta_1 \cap \theta_2 \Rightarrow s(A) = 1/2$

$A = \theta_1 \cup \theta_2 \cup \theta_3 = (\theta_1 \cup \theta_2) \cup \theta_3 = \theta_1 \cup (\theta_2 \cup \theta_3) = \theta_2 \cup (\theta_1 \cup \theta_3) \Rightarrow s(A) = 3$

$A = \theta_1 \cap \theta_2 \cap \theta_3 = (\theta_1 \cap \theta_2) \cap \theta_3 = \theta_1 \cap (\theta_2 \cap \theta_3) = \theta_2 \cap (\theta_1 \cap \theta_3) \Rightarrow s(A) = 1/3$

$A = (\theta_1 \cap \theta_2) \cup \theta_3 \Rightarrow s(A) = 3/2$

$A = (\theta_1 \cup \theta_2) \cap \theta_3 \Rightarrow s(A) = 2/3$

$A = (\theta_1 \cap \theta_2) \cup (\theta_3 \cap \theta_4) \Rightarrow s(A) = 1$

$A = (\theta_1 \cup \theta_2) \cap (\theta_3 \cup \theta_4) \Rightarrow s(A) = 1$

$A = (\theta_1 \cap \theta_2) \cup (\theta_3 \cap \theta_4 \cap \theta_5) \Rightarrow s(A) = 5/6$

$A = (\theta_1 \cup \theta_2) \cap (\theta_3 \cup \theta_4 \cup \theta_5) \Rightarrow s(A) = 6/5$

Thus the evaluation of $s(A)$ for any general irreductible proposition A can always be obtained from the two basic rules (46) and (47). This generalized definition makes sense with the notion of entropy and is coherent with classical definition (i.e. $H_g(m) \equiv H(m)$ when $m(.)$ becomes a bayesian bpa $p(.)$).

Proposition 5. *Let $\Theta = \{\theta_1, \ldots, \theta_n\}$ be a general frame of discernment of the problem under consideration and a general body of evidence with information granule $m(.)$ on D^Θ, then the generalized entropy $H_g(m)$ takes its minimal value $-n \ln(n)$ when the source provides the maximum of paradoxe which is obtained when $m(\theta_1 \cap \ldots \cap \theta_n) = 1$.*

But it is important to note that the maximum of uncertainty is not obtained when $m(\theta_1 \cup \ldots \cup \theta_n) = 1$ but rather for a specific bba $m(.)$ which distributes some weight of evidence assignment to each proposition $A \in D^\Theta$ because there is less information (from the information theory viewpoint) when there exists several propositions with non nul bba rather than there is only one. One has also to take into account the intrinsic self-information of the propositions to get a good measure of global information provided by a source. The generalized entropy includes both aspects of the information (the intrinsic and the classical aspect). The uniform distribution for $m(.)$ does not generate the maximum generalized-entropy because of the different intrinsic self-information of each proposition (see next example). The generalized entropy $H_g(m)$ of any source, defined with respect to a frame Θ with a given bba $m(.)$, appears to be a very promising and useful tool to measure the degree of uncertainty and paradoxe of any given source of information.

Example 14. We give here some values of $H_g(m)$ for different kinds of sources of information over the same frame $\Theta = \{\theta_1, \theta_2\}$. The sources have been classified from the most informative one $\mathcal{B}_1$ up to the less informative one $\mathcal{B}_{16}$. $\mathcal{B}_{16}$ corresponds to the source containing minimal information on the hyper-power set of the frame Θ (thus $\mathcal{B}_{16}$ has the minimal discrimination power between all possible propositions). There does not exist a source $\mathcal{B}_k$ such that $H_g^{\mathcal{B}_k}(m) > H_g^{\mathcal{B}_{16}}(m)$ for this simple example. Finding $m^*(.)$ such that $H_g(m^*)$ takes its maximal value for a general frame Θ with $|\Theta| = n$ is called the *general whitening source problem*. No solution for now for this problem has been obtained so far.

$\mathcal{B}_1$ is the most informative source because all the weights of evidence about the truth are focused only on the smaller element $\theta_1 \cap \theta_2$ of hyper-powerset D^Θ. $\mathcal{B}_2$ is less informative than $\mathcal{B}_1$ because there exists an ambiguity between the two propositions $\theta_1 \cup \theta_2$ and $\theta_1 \cap \theta_2$. $\mathcal{B}_3$ and $\mathcal{B}_4$ are less informative than $\mathcal{B}_1$ because the weights of evidence about the truth are focused on larger elements (θ_1 or θ_2 respectively) of D^Θ. $\mathcal{B}_6$ is less informative than $\mathcal{B}_3$ or $\mathcal{B}_4$ because the weight of evidence about the truth is focused on a bigger element $\theta_1 \cup \theta_2$ of D^Θ. $\mathcal{B}_7$ is less informative than previous sources since there exists an ambiguity between the two propositions θ_1 and θ_2 but it is more informative than $\mathcal{B}_9$ since

the discrimination power (our easiness to decide which proposition supports the truth) is higher with $\mathcal{B}_7$ than with $\mathcal{B}_9$. Note that even if in this very simple example, it is not obvious to see that $\mathcal{B}_{16}$ is the white source (less informative) of information. Most of readers would have probably thought to choose either $\mathcal{B}_6$ or $\mathcal{B}_{15}$. This comes from the confusion between the intrinsic information supported by the proposition itself and the information supported by the whole bba $m(.)$.

	$m(\theta_1)$	$m(\theta_2)$	$m(\theta_1 \cup \theta_2)$	$m(\theta_1 \cap \theta_2)$	$H_g(m)$
$\mathcal{B}_1$	0	0	0	1	-1.386
$\mathcal{B}_2$	0	0	0.3	0.7	-0.186
$\mathcal{B}_3$	1	0	0	0	0
$\mathcal{B}_4$	0	1	0	0	0
$\mathcal{B}_5$	0.1	0.2	0	0.7	0.081
$\mathcal{B}_6$	0	0	1	0	0.346
$\mathcal{B}_7$	0.8	0.2	0	0	0.500
$\mathcal{B}_8$	0	0	0.7	0.3	0.673
$\mathcal{B}_9$	0.5	0.5	0	0	0.693
$\mathcal{B}_{10}$	0.7	0.2	0.1	0	0.721
$\mathcal{B}_{11}$	0.7	0.2	0	0.1	0.893
$\mathcal{B}_{12}$	0.1	0.2	0.7	1	0.919
$\mathcal{B}_{13}$	0.1	0.2	0.3	0.4	1.015
$\mathcal{B}_{14}$	0.1	0.2	0.4	0.3	1.180
$\mathcal{B}_{15}$	0.25	0.25	0.25	0.25	1.299
$\mathcal{B}_{16}$	0.25	0.25	0.35	0.15	1.359

3.8 Blackman's example revisited

Let's take back the Blackman's example described in example 4 for the very simple assignment problem. In the DSmT framework, one has to deal with the following prior (predicted) and observed gbba defined on hyper-power set $D^\Theta = \{\emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2, \theta_1 \cap \theta_2\}$ as follows:

$$m_{T_1}(\theta_1) = 0.5 \quad m_{T_1}(\theta_2) = 0.5 \quad m_{T_1}(\theta_1 \cup \theta_2) = 0 \quad m_{T_1}(\theta_1 \cap \theta_2) = 0$$
$$m_{T_2}(\theta_1) = 0.1 \quad m_{T_2}(\theta_2) = 0.1 \quad m_{T_2}(\theta_1 \cup \theta_2) = 0.8 \quad m_{T_2}(\theta_1 \cap \theta_2) = 0$$
$$m_Z(\theta_1) = 0.5 \quad m_Z(\theta_2) = 0.5 \quad m_Z(\theta_1 \cup \theta_2) = 0 \quad m_Z(\theta_1 \cap \theta_2) = 0$$

Using the DSm rule of combination, we get now easily the following results

$$m_{T_1 Z}(\theta_1) = 0.25 \quad m_{T_1 Z}(\theta_2) = 0.25 \quad m_{T_1 Z}(\theta_1 \cap \theta_2) = 0.5$$
$$m_{T_2 Z}(\theta_1) = 0.45 \quad m_{T_2 Z}(\theta_2) = 0.45 \quad m_{T_2 Z}(\theta_1 \cap \theta_2) = 0.1$$

The values of the generalized entropy of the updated gbba $m_{T_1 Z}$ and $m_{T_2 Z}$ are $H_g(m_{T_1 Z}) \simeq 0.69$ and $H_g(m_{T_2 Z}) \simeq 1.04$. The increase of the generalized

entropies (i.e. the difference between the predicted and updated generalized entropies) are given by $\Delta_1 \triangleq H_g(m_{T_1 Z}) - H_g(m_{T_1}) = 0.69 - 0.69 = 0$ and $\Delta_2 \triangleq H_g(m_{T_2 Z}) - H_g(m_{T_2}) = 1.04 - 0.83 = 0.21$. This result means that the incorrect assignment $m_{T_2} \oplus m_Z$ has noticeably increased the generalized entropy of the system as one would have rightfully expected. The best assignment solution is obtained by selecting the fusion (assignment between a track T and a measurement Z) which generates the smallest increase of the generalized entropy. In this framework and in this case, the Tchamova's approach based on the minimum city-block or euclidean distances provides also the correct assignment Z with track T_1 since $d_1(T_1, T_1 Z) < d_1(T_2, T_2 Z)$ and $d_2(T_1, T_1 Z) < d_2(T_2, T_2 Z)$ because one has

$$d_1(T_1, T_1 Z) = 1 \qquad \text{and} \qquad d_1(T_2, T_2 Z) = 1.6$$
$$d_2(T_1, T_1 Z) = 0.612 \qquad \text{and} \qquad d_2(T_2, T_2 Z) = 0.946$$

Neither the use of classical entropy $H(m)$ nor the entropy evaluated from pignistic probabilities allow us to get the correct assignment solution from the DST framework in this example .

Let's consider now the previous predicted gbba $m_{T_1}(.)$ and $m_{T_2}(.)$ but now with an observation bba which agrees with $m_{T_2}(.)$, i.e.

$$m_Z(\theta_1) = 0.1 \qquad m_Z(\theta_2) = 0.1 \qquad m_Z(\theta_1 \cup \theta_2) = 0.8 \qquad m_Z(\theta_1 \cap \theta_2) = 0$$

Using the DSm rule of combination, we get now the following results

$$m_{T_1 Z}(\theta_1) = 0.45 \; m_{T_1 Z}(\theta_2) = 0.45 \; m_{T_1 Z}(\theta_1 \cup \theta_2) = 0 \qquad m_{T_1 Z}(\theta_1 \cap \theta_2) = 0.1$$
$$m_{T_2 Z}(\theta_1) = 0.17 \; m_{T_2 Z}(\theta_2) = 0.17 \; m_{T_1 Z}(\theta_1 \cup \theta_2) = 0.64 \; m_{T_2 Z}(\theta_1 \cap \theta_2) = 0.02$$

The generalized entropies of the two possible assignments take now the following values $H_g(m_{T_1 Z}) \simeq 1.0405$ and $H_g(m_{T_2 Z}) \simeq 1.0958$ which are very close but the entropy increases become now $\Delta_1 \triangleq H_g(m_{T_1 Z}) - H_g(m_{T_1}) = 1.0405 - 0.69 = 0.35$ and $\Delta_2 \triangleq H_g(m_{T_2 Z}) - H_g(m_{T_2}) = 1.095 - 0.83 = 0.265$. By selecting the smallest increase of the generalized entropies, we get again the correct assignment Z with track T_2 for this second case. As within the same example discussed in the DST framework, the minimum distance approach fails here to obtain the correct assignment since one has now $d_1(T_1, T_1 Z) < d_1(T_2, T_2 Z)$ and $d_2(T_1, T_1 Z) < d_2(T_2, T_2 Z)$ because

$$d_1(T_1, T_1 Z) = 0.2 \qquad \text{and} \qquad d_1(T_2, T_2 Z) = 0.32$$
$$d_2(T_1, T_1 Z) = 0.122 \qquad \text{and} \qquad d_2(T_2, T_2 Z) = 0.189$$

In concluding remark, we have shown through this simple example, how a simple and unique criteria based on our *generalized entropy like measure* drawn from our DSmT can serve as an useful tool to solve the assignment problem for both cases investigated here. No case-dependent approach is then required here to get the correct solution as we had already argued in example 4. However, more theoretical investigations must be done to prove that our criteria is actually the best one to solve the assignment problem in general.

3.9   Zadeh's example revisited

Let's take back the disturbing Zadeh's example [74] given in section 2.4. Two doctors examine a patient and agree that it suffers from either meningitis (M), concussion (C) or brain tumor (T). Thus $\Theta = \{M, C, T\}$. Assume that the two doctors agree in their low expectation of a tumor, but disagree in likely cause and provide the following diagnosis

$$m_1(M) = 0.99 \qquad m_1(T) = 0.01$$

and $\forall A \in D^{\Theta}, A \neq T, A \neq M, m_1(A) = 0$

$$m_2(C) = 0.99 \qquad m_2(T) = 0.01$$

and $\forall A \in D^{\Theta}, A \neq T, A \neq C, m_2(A) = 0$
The new general rule of combination (26), yields the following combined information granule

$$m(M \cap C) = 0.9801 \qquad m(M \cap T) = 0.0099$$
$$m(C \cap T) = 0.0099 \qquad m(T) = 0.0001$$

From this granule, one gets

$$\text{Bel}(M) = m(M \cap C) + m(M \cap T) = 0.99$$
$$\text{Bel}(C) = m(M \cap C) + m(T \cap C) = 0.99$$
$$\text{Bel}(T) = m(T) + m(M \cap T) + m(C \cap T) = 0.0199$$

If both doctors can be considered as equally reliable, the combined information granule $m(.)$ mainly focuses weight of evidence on the paradoxical proposition $M \cap C$ which means that patient suffers both meningitis and concussion but almost surely not from brain tumor. This conclusion is coherent with the common sense actually. Then, no therapy for brain tumor (like heavy and ever risky brain surgical intervention) will be chosen in such case. This really helps to take important decision to save the life of the patient in this example. A deeper medical examination adapted to both meningitis and concussion will almost surely be done before applying the best therapy for the patient. Just remember that in this case, the DST had concluded that the patient had brain tumor with certainty . . . .

3.10   Mahler's example revisited

Let's consider now the following example excerpt from the Ronald Mahler's paper [36]. We consider that our classification knowledge base consists of the three (imaginary) new and rare diseases corresponding to following frame of discernment

$$\Theta = \{\theta_1 = kotosis, \theta_2 = phlegaria, \theta_3 = pinpox\}$$

We assume that the three diseases are equally likely to occur in the patient population but there is some evidence that *phlegaria* and *pinpox* are the same disease and there is also a small possibility that *kotosis* and *phlegaria* might be the same disease. Finally, there is a small possibility that all three diseases are the same. This information can be expressed by assigning a priori bba as follows

$$m_0(\theta_1) = 0.2 \qquad m_0(\theta_2) = 0.2 \qquad m_0(\theta_3) = 0.2$$
$$m_0(\theta_2 \cap \theta_3) = 0.2 \; m_0(\theta_1 \cap \theta_2) = 0.1 \; m_0(\theta_1 \cap \theta_2 \cap \theta_3) = 0.1$$

Let Bel(.) the prior belief measure corresponding to this prior bba $m(.)$. Now assume that Doctor D_1 and Doctor D_2 examine a patient and deliver diagnoses with following reports:

– Report for D_1: $m_1(\theta_1 \cup \theta_2 \cup \theta_3) = 0.05$ $m_1(\theta_2 \cup \theta_3) = 0.95$
– Report for D_2: $m_2(\theta_1 \cup \theta_2 \cup \theta_3) = 0.20$ $m_2(\theta_2) = 0.80$

The combination of the evidences provided by the two doctors $m' = m_1 \oplus m_2$ obtained by the general rule of combination (26) yields the following bba $m'(.)$

$$m'(\theta_2) = 0.8 \qquad m'(\theta_2 \cup \theta_3) = 0.19 \qquad m'(\theta_1 \cup \theta_2 \cup \theta_3) = 0.01$$

The combination of bba $m'(.)$ with prior evidence $m_0(.)$ yields the final bba $m = m_0 \oplus m' = m_0 \oplus [m_1 \oplus m_2]$ with

$$m(\theta_1) = 0.002 \qquad\qquad m(\theta_2) = 0.200 \qquad m(\theta_3) = 0.040$$
$$m(\theta_1 \cap \theta_2) = 0.260 \qquad\quad m(\theta_2 \cap \theta_3) = 0.360 \; m(\theta_1 \cap \theta_2 \cap \theta_3) = 0.100$$
$$m(\theta_1 \cap (\theta_2 \cup \theta_3)) = 0.038$$

Therefore the final belief function given by (22) is

$$\text{Bel}(\theta_1) = 0.002 + 0.260 + 0.100 + 0.038 = 0.400$$
$$\text{Bel}(\theta_2) = 0.200 + 0.260 + 0.360 + 0.100 = 0.920$$
$$\text{Bel}(\theta_3) = 0.040 + 0.360 + 0.100 = 0.500$$
$$\text{Bel}(\theta_1 \cap \theta_2) = 0.260 + 0.100 = 0.360$$
$$\text{Bel}(\theta_2 \cap \theta_3) = 0.360 + 0.100 = 0.460$$
$$\text{Bel}(\theta_1 \cap (\theta_2 \cup \theta_3)) = 0.038 + 0.100 = 0.138$$
$$\text{Bel}(\theta_1 \cap \theta_2 \cap \theta_3) = 0.100$$

Thus, on the basis of all the evidences one has, we are able to conclude with high a degree of belief that the patient has phlegaria which is coherent with the Mahler's conclusion based on his Conditioned Dempster-Shafer theory developed from his conditional event algebra although a totally new and simpliest approach has been adopted here.

3.11   A thief identification example

Let's revisit now a very simple and classical thief identification example. Assume that a 75 years old grandfather is taking a walk with his 9 years old grandson in a park. They saw at 50 meters away, a 45 years old pickpocket robbering the bag of an old lady. A policeman looking for some witnesses of this event asks separately the grandfather and his grandchild if they have seen the thief (they both answer yes) and how old approximately was the thief (a young or an old man). The grandfather (source of information B_1 reports that the thief was a young man with high confidence 0.99 and with only a low uncertainty 0.01. His grandson reports that the thief was a old man with high confidence 0.99 and with only a low uncertainty 0.01. These two witnesses provide fair reports (with respect to their own world of knowledge) even if apparently they appear as almost fully paradoxical. The policeman then send the two reports with only the minimal information about witnesses (saying only their names and that they were a priori fully trustable) to an investigator. The investigator has no possibility to meet or to call back the witnesses in order to get more details.

Under such condition, what would be the best reasoning of the investigator to infer the age of the thief to eventually help to catch him? Such kind of simple example occurs quite frequently in many witnesses problems actually. A rational investigator will almost surely suspect a mistake or an error in one or both reports since they appear apparently in (almost) full contradiction. The investigator will then try to take his final decision with some other better information (if any). If the investigator uses our new plausible and paradoxical reasoning, he will defined the following bba with respect to the frame of discernment $\Theta = \{\theta_1 = \text{young}, \theta_2 = \text{old}\}$ and the available reports $\mathcal{B}_1$ and $\mathcal{B}_2$ with bba

$$m_1(\theta_1) = 0.99 \qquad m_1(\theta_2) = 0 \qquad m_1(\theta_1 \cup \theta_2) = 0.01 \qquad m_1(\theta_1 \cap \theta_2) = 0$$

$$m_2(\theta_1) = 0 \qquad m_2(\theta_2) = 0.99 \qquad m_2(\theta_1 \cup \theta_2) = 0.01 \qquad m_2(\theta_1 \cap \theta_2) = 0$$

The fusion of these two sources of information yields the global bba $m(.)$ with

$$m(\theta_1) = 0.0099 \qquad m(\theta_2) = 0.0099$$

$$m(\theta_1 \cup \theta_2) = 0.0001 \qquad m(\theta_1 \cap \theta_2) = 0.9801$$

Thus, from this global information, the investigator has no better choice but to consider with almost certainty that the thief was both a young and old man. By assuming that the expected life duration is around 80 years, the inspector will deduce that the true age of the thief is around 40 years old which is not too far from the truth. At least, this conclusion could be helpful to interrogate some suspicious individuals.

3.12   A model to generate information granules $m(.)$ from intervals

We present here a model to generate information granules $m(.)$ from information represented by intervals. It is very common in practice that uncertain sources of information provide evidence on a given proposition in term of basic intervals $[\epsilon_*, \epsilon^*] \subset [0, 1]$ rather than a direct bba $m(.)$. In such cases, some preprocessing must be done before applying the general rule of combination between such sources to take the final decision.

In the DST framework, we recall that the simpliest and easiest transformation to convert $[\epsilon_*, \epsilon^*]$ into bba has already been proposed by A. Appriou in [1] and successfully used in [14]. The basic idea was to interpret ϵ_* as the minimal credibility committed to A and ϵ^* as the plausibility committed to A. In other words, the Appriou's transformation model within the DST framework is the following one

$$\epsilon_* = m(A)$$
$$\epsilon^* = 1 - m(A^c)$$
$$\epsilon^* - \epsilon_* = m(A \cup A^c)$$

This model can be directly extended within our new theory of plausible and paradoxical reasoning by setting now[4]

$$\epsilon_* = m(A) + \frac{1}{2}m(A \cap A^c)$$
$$\epsilon^* = 1 - m(A^c) - \frac{1}{2}m(A \cap A^c)$$
$$\epsilon^* - \epsilon_* = m(A \cup A^c)$$

or equivalently

$$m(A) + \frac{1}{2}m(A \cap A^c) = \epsilon_* \tag{48}$$

$$m(A^c) + \frac{1}{2}m(A \cap A^c) = 1 - \epsilon^* \tag{49}$$

$$m(A \cup A^c) = \epsilon^* - \epsilon_* \tag{50}$$

This appealing model presents nice properties specially when $\epsilon^* = \epsilon_* = 0$ or when $\epsilon^* = \epsilon_* = 1$. This model is moreover coherent with the previous Appriou's model whenever the source becomes rational (i.e $m(A \cap A^c) = 0$). This new model presents however a degree of freedom since one has only two constraints (48) and (49) for three unknowns $m(A)$, $m(A^c)$ and $m(A \cap A^c)$. Thus in general,

---

[4] The notation A^c has been kept here for simplicity but in our DSmT A^c must not be interpreted directly as the complement of A since $m(A \cap A^c)$ can take a positive value ≤ 1 but as a (partial overlapping) *paradoxical alternative* (see the forthcoming numerical examples).

without an additional constraint, there exists many possible choices for $m(A)$, $m(A^c)$ and $m(A \cap A^c)$ and, therefore, several bba $m(.)$ satisfying this transformation model. Without extra prior information, it becomes difficult to justify the choice of a specific bba versus all other admissible possibilities for $m(.)$.

To solve this important drawback, we propose to add the constraint on the maximization of the generalized-entropy $H_g(m)$. This will allow us to obtain from $[\epsilon_*, \epsilon^*]$ the unique bba $m(.)$ having the minimum of specificity and admissible with our transformation model. From definition of $H_g(m)$ and previous equations (48)-(50), one gets

$$\begin{aligned}
H_g(m) = &- (\epsilon_* - m(A \cap A^c)/2)\ln(\epsilon_* - m(A \cap A^c)/2) \\
&- (1 - \epsilon^* - m(A \cap A^c)/2)\ln(1 - \epsilon^* - m(A \cap A^c)/2) \\
&- \frac{1}{2}(\epsilon^* - \epsilon_*)\ln(\frac{1}{2}(\epsilon^* - \epsilon_*)) \\
&- 2m(A \cap A^c)\ln(2m(A \cap A^c))
\end{aligned}$$

The maximization of $H_g(m)$ is obtained for the optimal value $m^\star(A \cap A^c)$ such that $\frac{\partial H_g}{\partial m(A \cap A^c)}(m^\star(A \cap A^c)) = 0$ and $\frac{\partial^2 H_g}{\partial m(A \cap A^c)^2}(m^\star(A \cap A^c)) < 0$. The annulation of the first derivative is obtained by the solution of the equation

$$\frac{1}{2}\ln(\epsilon_* - m^\star/2) + \frac{1}{2}\ln(1 - \epsilon^* - m^\star/2) - 2m^\star\ln(2m^\star) - 1 = 0$$

or equivalently after basic algebraic manipulations

$$64e^2(m^\star)^4 - (m^\star)^2 + 2(1 - \epsilon^* + \epsilon_*)m^\star - 4(1 - \epsilon^*)\epsilon_* = 0 \tag{51}$$

The solution of this equation can be easily found using classical numerical methods. It is also easy to check that the second derivative is always negative and therefore $H_g(m)$ reaches its maximal value when

$$m(A) + \frac{1}{2}m^\star(A \cap A^c) = \epsilon_* \tag{52}$$

$$m(A^c) + \frac{1}{2}m^\star(A \cap A^c) = 1 - \epsilon^* \tag{53}$$

$$m(A \cup A^c) = \epsilon^* - \epsilon_* \tag{54}$$

This completes the definition of our new transformation model. Note that $[\epsilon_*, \epsilon^*]$ can also be generated from bba $m(.)$ through (48)-(50).

Example 15. for $[\epsilon_*, \epsilon^*] = [0.0, 0.0]$, one gets

$m(A \cap A^c) = 0.000 \qquad m(A) = 0.000 \qquad m(A^c) = 1.000 \qquad m(A \cup A^c) = 0.000$

Example 16. for $[\epsilon_*, \epsilon^*] = [0.2, 0.2]$, one gets

$m(A \cap A^c) \approx 0.164 \qquad m(A) \approx 0.118 \qquad m(A^c) \approx 0.718 \qquad m(A \cup A^c) = 0.000$

Example 17. for $[\epsilon_*, \epsilon^*] = [0.5, 0.5]$, one gets

$m(A \cap A^c) \approx 0.192$ $m(A) \approx 0.404$ $m(A^c) \approx 0.404$ $m(A \cup A^c) = 0.000$

Example 18. for $[\epsilon_*, \epsilon^*] = [0.8, 0.8]$, one gets

$m(A \cap A^c) \approx 0.164$ $m(A) \approx 0.718$ $m(A^c) \approx 0.118$ $m(A \cup A^c) = 0.000$

Example 19. for $[\epsilon_*, \epsilon^*] = [1.0, 1.0]$, one gets

$m(A \cap A^c) = 0.000$ $m(A) = 1.000$ $m(A^c) = 0.000$ $m(A \cup A^c) = 0.000$

Example 20. for $[\epsilon_*, \epsilon^*] = [0.2, 0.4]$, one gets

$m(A \cap A^c) \approx 0.152$ $m(A) \approx 0.124$ $m(A^c) \approx 0.524$ $m(A \cup A^c) = 0.200$

Example 21. for $[\epsilon_*, \epsilon^*] = [0.6, 0.8]$, one gets

$m(A \cap A^c) \approx 0.152$ $m(A) \approx 0.524$ $m(A^c) \approx 0.124$ $m(A \cup A^c) = 0.200$

Example 22. for $[\epsilon_*, \epsilon^*] = [0.4, 0.6]$, one gets

$m(A \cap A^c) \approx 0.170$ $m(A) \approx 0.315$ $m(A^c) \approx 0.315$ $m(A \cup A^c) = 0.200$

Example 23. for $[\epsilon_*, \epsilon^*] = [0.3, 0.9]$, one gets

$m(A \cap A^c) \approx 0.100$ $m(A) \approx 0.250$ $m(A^c) \approx 0.050$ $m(A \cup A^c) = 0.600$

Example 24. for $[\epsilon_*, \epsilon^*] = [0.0, 1.0]$, one gets

$m(A \cap A^c) = 0.000$ $m(A) = 0.000$ $m(A^c) = 0.000$ $m(A \cup A^c) = 1.000$

4   Conclusion

In this paper, the foundations for a new theory of paradoxical and plausible reasoning have been developed. The DSmT takes into account in the combination process itself the possibility for uncertain and paradoxical information. The basis for the development of this theory is to work with the hyper-power set of the frame of discernment relative to the problem under consideration rather than its classical power set since, in general, the frame of discernment cannot be fully described in terms of an exhaustive and exclusive list of disjoint elementary hypotheses. In such general case, no refinement is possible to apply directly the Dempster-Shafer theory (DST) of evidence. In DSmT, the rule of combination is justified from the maximum entropy principle and there is no mathematical impossibility to combine sources of evidence even if they appear at first glance in contradiction (in the Shafer's sense) since the paradox between sources is fully taken into account in our formalism. We have also shown that in general, the combination of evidence yields unavoidable paradoxes. This theory has shown, through many illustrated examples, that conclusion drawn from it, can provide a result which agrees with the human reasoning and can be very helpful to take a decision for some complex problems where the classical DST usually fails. This new theory provides also a theoretical bridge between the combination of paradoxical source of information and the Smarandache's logic.

Appendix

We prove here that the hyper-power set D^Θ of $\Theta = \{\theta_1, \theta_2, \theta_3\}$, is given by the set of the following 19 irreducible propositions:

$$\alpha_0 \triangleq \emptyset$$

$$\alpha_1 \triangleq \theta_1 \qquad\qquad \alpha_{10} \triangleq \theta_1 \cup \theta_2 \cup \theta_3$$
$$\alpha_2 \triangleq \theta_2 \qquad\qquad \alpha_{11} \triangleq \theta_1 \cap \theta_2 \cap \theta_3$$
$$\alpha_3 \triangleq \theta_3 \qquad\qquad \alpha_{12} \triangleq (\theta_1 \cup \theta_2) \cap \theta_3$$
$$\alpha_4 \triangleq \theta_1 \cup \theta_2 \qquad\qquad \alpha_{13} \triangleq (\theta_1 \cup \theta_3) \cap \theta_2$$
$$\alpha_5 \triangleq \theta_1 \cup \theta_3 \qquad\qquad \alpha_{14} \triangleq (\theta_2 \cup \theta_3) \cap \theta_1$$
$$\alpha_6 \triangleq \theta_2 \cup \theta_3 \qquad\qquad \alpha_{15} \triangleq (\theta_1 \cap \theta_2) \cup \theta_3$$
$$\alpha_7 \triangleq \theta_1 \cap \theta_2 \qquad\qquad \alpha_{16} \triangleq (\theta_1 \cap \theta_3) \cup \theta_2$$
$$\alpha_8 \triangleq \theta_1 \cap \theta_3 \qquad\qquad \alpha_{17} \triangleq (\theta_2 \cap \theta_3) \cup \theta_1$$
$$\alpha_9 \triangleq \theta_2 \cap \theta_3 \qquad\qquad \alpha_{18} \triangleq (\theta_1 \cup \theta_2) \cap (\theta_1 \cup \theta_3) \cap (\theta_2 \cup \theta_3)$$

We need to verify that $\forall \alpha_i \in D^\Theta, \forall \alpha_j \in D^\Theta, (\alpha_i \cup \alpha_j) \in D^\Theta$ and $(\alpha_i \cap \alpha_j) \in D^\Theta$.

First note that $\forall \alpha_i, i = 0, \ldots 18$, one always has

$$\alpha_0 \cap \alpha_i = \alpha_0 \qquad \text{and} \qquad \alpha_0 \cup \alpha_i = \alpha_i$$

Let's compute now all $\alpha_i \cap \alpha_j$ for $i, j = 1, \ldots, 18$. Using classical intersection operator on sets, we get the following result summarized in the following symmetric table

$\cap$	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}	α_{16}	α_{17}	α_{18}
α_1	α_1	α_7	α_8	α_1	α_1	α_{14}	α_7	α_8	α_{11}	α_1	α_{11}	α_8	α_7	α_{14}	α_{14}	α_{14}	α_1	α_{14}
α_2	α_7	α_2	α_9	α_2	α_{13}	α_2	α_7	α_{11}	α_9	α_2	α_{11}	α_9	α_{13}	α_7	α_{13}	α_2	α_7	α_{13}
α_3	α_8	α_9	α_3	α_{12}	α_3	α_3	α_{11}	α_8	α_9	α_3	α_{11}	α_{12}	α_9	α_8	α_3	α_{12}	α_{12}	α_{12}
α_4	α_1	α_2	α_{12}	α_4	α_{17}	α_{16}	α_7	α_8	α_9	α_4	α_{11}	α_{12}	α_{13}	α_{14}	α_{18}	α_{16}	α_{17}	α_{18}
α_5	α_1	α_{13}	α_3	α_{17}	α_5	α_{15}	α_7	α_8	α_9	α_9	α_{11}	α_{12}	α_{13}	α_{14}	α_{18}	α_{18}	α_{17}	α_{18}
α_6	α_{14}	α_2	α_3	α_{16}	α_{15}	α_6	α_7	α_8	α_9	α_6	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}	α_{16}	α_{18}	α_{18}
α_7	α_7	α_7	α_{11}	α_7	α_7	α_7	α_7	α_{11}	α_{11}	α_7	α_{11}	α_{12}	α_7	α_7	α_7	α_{16}	α_7	α_7
α_8	α_8	α_{11}	α_8	α_8	α_8	α_8	α_{11}	α_8	α_{11}	α_8	α_{11}	α_{12}	α_{11}	α_8	α_8	α_8	α_8	α_8
α_9	α_{11}	α_9	α_9	α_9	α_9	α_9	α_{11}	α_{11}	α_9	α_9	α_{11}	α_9	α_9	α_{11}	α_9	α_9	α_9	α_9
α_{10}	α_1	α_2	α_3	α_4	α_9	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}	α_{16}	α_{17}	α_{18}
α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}	α_{11}
α_{12}	α_8	α_9	α_{12}	α_{12}	α_{12}	α_{12}	α_{12}	α_{12}	α_9	α_{12}	α_{11}	α_{12}	α_{11}	α_8	α_{12}	α_{12}	α_{12}	α_{12}
α_{13}	α_7	α_{13}	α_9	α_{13}	α_{13}	α_{13}	α_7	α_{11}	α_9	α_{13}	α_{11}	α_{11}	α_{13}	α_7	α_{13}	α_{13}	α_{13}	α_{13}
α_{14}	α_{14}	α_7	α_8	α_{14}	α_{14}	α_{14}	α_7	α_8	α_{11}	α_{14}	α_{11}	α_8	α_7	α_{14}	α_{14}	α_{14}	α_{14}	α_{14}
α_{15}	α_{14}	α_{13}	α_3	α_{18}	α_{18}	α_{15}	α_7	α_8	α_9	α_{15}	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}	α_{18}	α_{18}	α_{18}
α_{16}	α_{14}	α_2	α_{12}	α_{16}	α_{18}	α_{16}	α_{16}	α_8	α_9	α_{16}	α_{11}	α_{12}	α_{13}	α_{14}	α_{18}	α_{16}	α_{18}	α_{18}
α_{17}	α_1	α_7	α_{12}	α_{17}	α_{17}	α_{18}	α_7	α_8	α_9	α_{17}	α_{11}	α_{12}	α_{13}	α_{14}	α_{18}	α_{18}	α_{17}	α_{18}
α_{18}	α_{14}	α_{13}	α_{12}	α_{18}	α_{18}	α_{18}	α_7	α_8	α_9	α_{18}	α_{11}	α_{12}	α_{13}	α_{14}	α_{18}	α_{18}	α_{18}	α_{18}

Hence we have just proved here that $\forall \alpha_i, \alpha_j \in D^\Theta$, $\alpha_i \cap \alpha_j \in D^\Theta$. It remains now to compute all $\alpha_i \cup \alpha_j$ for $i, j = 1, \ldots, 18$. Using classical union operator on sets, we get the following result summarized in the following symmetric table

$\cup$	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}	α_{16}	α_{17}	α_{18}
α_1	α_1	α_4	α_5	α_4	α_5	α_{10}	α_1	α_1	α_{17}	α_{10}	α_1	α_{17}	α_{17}	α_1	α_5	α_4	α_{17}	α_{17}
α_2	α_4	α_2	α_6	α_4	α_{10}	α_6	α_2	α_{16}	α_2	α_{10}	α_2	α_{16}	α_2	α_{16}	α_6	α_{16}	α_4	α_{16}
α_3	α_5	α_6	α_3	α_{10}	α_5	α_6	α_{15}	α_3	α_3	α_{10}	α_3	α_3	α_{15}	α_{15}	α_{15}	α_6	α_5	α_{15}
α_4	α_4	α_4	α_{10}	α_4	α_{10}	α_{10}	α_4	α_4	α_4	α_{10}	α_4	α_4	α_4	α_4	α_{10}	α_{10}	α_{10}	α_{10}
α_5	α_5	α_{10}	α_5	α_{10}	α_5	α_{10}	α_5	α_5	α_5	α_{10}	α_5	α_5	α_5	α_5	α_5	α_{10}	α_5	α_5
α_6	α_{10}	α_6	α_6	α_{10}	α_{10}	α_6	α_6	α_6	α_6	α_{10}	α_6	α_6	α_6	α_6	α_6	α_6	α_{10}	α_6
α_7	α_1	α_2	α_{15}	α_4	α_5	α_6	α_7	α_{14}	α_{13}	α_{10}	α_7	α_{18}	α_{13}	α_{14}	α_{15}	α_{16}	α_{17}	α_{18}
α_8	α_1	α_{16}	α_3	α_4	α_5	α_6	α_{14}	α_8	α_{12}	α_{10}	α_8	α_{12}	α_{18}	α_{14}	α_{15}	α_{16}	α_{17}	α_{18}
α_9	α_{17}	α_2	α_3	α_4	α_5	α_6	α_{13}	α_{12}	α_9	α_{10}	α_9	α_{12}	α_{13}	α_{18}	α_{15}	α_{16}	α_{17}	α_{18}
α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}	α_{10}
α_{11}	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}	α_{16}	α_{17}	α_{18}
α_{12}	α_{17}	α_{16}	α_3	α_4	α_5	α_6	α_{18}	α_{12}	α_{12}	α_{10}	α_{12}	α_{12}	α_{18}	α_{18}	α_{15}	α_{16}	α_{17}	α_{18}
α_{13}	α_{17}	α_2	α_{15}	α_4	α_5	α_6	α_{13}	α_{18}	α_{13}	α_{10}	α_{13}	α_{18}	α_{13}	α_{18}	α_{15}	α_{16}	α_{17}	α_{18}
α_{14}	α_1	α_{16}	α_{15}	α_4	α_5	α_6	α_{14}	α_{14}	α_{18}	α_{10}	α_{14}	α_{18}	α_{18}	α_{14}	α_{15}	α_{16}	α_{17}	α_{18}
α_{15}	α_5	α_6	α_{15}	α_{10}	α_5	α_6	α_{15}	α_{15}	α_{15}	α_{10}	α_{15}	α_{15}	α_{15}	α_{15}	α_{15}	α_6	α_5	α_{15}
α_{16}	α_4	α_{16}	α_6	α_4	α_{10}	α_6	α_{16}	α_{16}	α_{16}	α_{10}	α_{16}	α_{16}	α_{16}	α_{16}	α_6	α_{16}	α_4	α_{16}
α_{17}	α_{17}	α_4	α_5	α_4	α_5	α_{10}	α_{17}	α_{17}	α_{17}	α_{10}	α_{17}	α_{17}	α_{17}	α_{17}	α_5	α_4	α_{17}	α_{17}
α_{18}	α_{17}	α_{16}	α_{15}	α_4	α_5	α_6	α_{18}	α_{18}	α_{18}	α_{10}	α_{18}	α_{18}	α_{18}	α_{18}	α_{15}	α_{16}	α_{17}	α_{18}

Therefore, one has proved that $\forall \alpha_i \in D^\Theta, \forall \alpha_j \in D^\Theta, (\alpha_i \cup \alpha_j) \in D^\Theta$ and $(\alpha_i \cap \alpha_j) \in D^\Theta$ and the set $\{\alpha_0, \ldots, \alpha_{18}\}$ corresponds effectively to the hyper-power set of $\Theta = \{\theta_1, \theta_2, \theta_3\}$ we were looking for.

References

1. Appriou A.,"Procédure d'aide à la décision multi-informateurs. Application à la classification multi-capteurs de cibles", AGARD Avionics Panel Symposium on Software Engineering and Its Applications, Turkey, April 25-29, 1988.

2. Ashbacher C.,"Introduction to Neutrosophic Logic", American research Press, 2002.

3. Bianaghi E., Madella P.,"Inductive and Deductive Reasoning techniques for Fuzzy Dempster-Shafer Classifiers", Proc. of 1997 Seventh IFSA World Congress, Prague, pp. 197-302, 1997.

4. Blackman S., Popoli R.,"Design and Analysis of Modern Tracking", 1232 pages, Artech House, 2000.

5. Cheng Y., Kashyap R.L., "Study of the Different Methods for Combining Evidence", Proceedings of SPIE on Applications of Artificial Intelligence, Vol. 635, pp. 384-393, 1986.

6. Comtet L., "Sperner Systems", Advanced Combinatorics: The Art of Finite and Infinite Expansions, Dordrecht, Netherlands:Reidel, pp. 271-273, 1974.

7. Cover T., Thomas J.A. "Elements of Information Theory", Wiley Series in Telecommunications, John Wiley and Sons, Inc., 1991.

8. De Brucq D., Colot O., Sombo A.,"Identical Foundation of Probability Theory and Fuzzy Set Theory", in Proc. of the 5th Intern. Conf. on Inf. Fusion (Fusion 2002), Annapolis, MD, USA, pp. 1442-1449, July 8-11, 2002.

9. Dedekind R.,"Über Zerlegungen von Zhalen durch ihre grössten gemeinsammen Teiler", In Gesammelte Werke, Bd. 1., pp. 103-148, 1897.

10. Dempster A.P.,"Upper and Lower Probabilities Induced by Multivalued Mapping", Annals of Mathematical Statistics, Vol. 28, pp. 325-339, 1967.

11. Dempster A.P.,"A Generalization of Bayesian Inference", Journal of the Royal Statistical Society, Serie B, Vol. 30, pp. 205-247, 1968.

12. Denoeux, T.,"Application of evidence theory to k-NN pattern classification", in Pattern Recognition in Practice IV, Gelsema, E. S. and Kanal, L. N. Editors, Elsevier, Amsterdam, pp. 13-24, 1994.

13. Dezert J., "Vers un nouveau concept de navigation autonome d'engin. Un lien entre le PDAF et la théorie de l'évidence", Ph. D. Thesis, University Paris XI, Orsay, France, Sept. 27, 1990.

14. Dezert J., "Autonomous Navigation with Uncertain Reference Points using the PDAF", Chapter 9 in Multitarget-Multisensor Tracking: Applications and Advances, Vol. 2, pp. 271-324,Y. Bar-Shalom Editor, Artech House, 1991.

15. Dezert J., "Optimal Bayesian Fusion of Multiple Unreliable Classifiers", Proceedings of 4th Intern. Conf. on Information Fusion (Fusion 2001), Montréal, Aug. 7-10, 2001.

16. Dezert J., "An introduction to the theory of plausible and paradoxical reasoning", Proceedings of NM&A 02, International Conference on Numerical methods and Applications, Springer Verlag Ed., Borovetz, Bulgaria, August 20-24, 2002.

17. Dubois D., Prade H.,"On the Unicity of Dempster Rule of Combination", International Journal of Intelligent Systems, Vol. 1, pp. 133-142, 1986.

18. Dubois D., Garbolino P., Kyburg H.E., Prade H., Smets, Ph.,"Quantified Uncertainty", J. Applied Non-Classical Logics, Vol.1, pp. 105-197, 1991.

19. Dubois D., Prade H.,"Evidence, Knowledge and Belief Functions", International Journal of Approximate Reasoning, Vol. 6, pp 295-319, 1992.

20. Erdös P., Ko C., Rado R.,"Intersection Theorems for Systems of Finite Sets", Quart. J. Math. Oxford, 12, pp. 313-320, 1961.

21. Guiasu S., Theodorescu R.,"La théorie mathématique de l'information", Dunod Editeur, Paris, 1968.

22. Hilton A.J., Milner E.C.,"Some Intersection Theorems of Systems of Finite Sets", Quart. J. Math. Oxford, 18, pp. 369-384, 1967.

23. Jaynes E.T.,"Where do we stand on Maximum Entropy ?", presented at the Maximum Entropy Formalism Conference, 104 pages, MIT, May 2-4, 1978.

24. Jaynes E.T.,"Probability Theory: The Logic of Science", Fragmentary Edition of March 1996, full edition in preparation by Cambridge Univ. Press, MA, USA.

25. Kennes R., Smets Ph., "Fast algorithms for Dempster-Shafer theory", in Uncertainty in Knowledge Bases, B. Bouchon-Meunier, R.R. Yager, L.A. Zadeh (Editors), Lecture Notes in Computer Science 521, Springer-Verlag, Berlin, pp. 14-23, 1991.

26. Kennes R., "Computational Aspects of the Möbius Transformation of Graphs", IEEE Trans. on SMC. 22, pp. 201-223, 1992.

27. Klawonn F., Smets Ph., "The dynamic of belief in the transferable belief model and specialization- generalization matrices", in Uncertainty in Artificial Intelligence 92, pp 130-137, Dubois D. and Wellman M. P. and D'Ambrosio B. and Smet, Ph. Editors, Morgan Kaufman, San Mateo, Ca, 1992.

28. Kleitman D., "On Dedekind's Problem: The Number of Monotone Boolean Functions", Proc. Amer. Math. Soc., Vol. 21, pp. 677-682, 1969.

29. Kleitman D., Markowsky G., "On Dedekind's Problem: The Number of Isotone Monotone Boolean Functions II", Trans. Amer. Math. Soc., Vol. 213, pp. 373-390, 1975.

30. Klopotek M.A., "Interpretation of belief function in Dempster-Shafer theory", Foundations of Computing and Decison Sciences, Vol. 20, no. 4, pp. 287-306, 1995.

31. Klopotek M.A., "Identification of Belief Structure in Dempster-Shafer Theory", Foundations of Computing and Decison Sciences, Vol. 21, no. 1, pp. 35-54, 1996.

32. Kohlas J., Monney P. A., "A Mathematical Theory of Hints: An Approach to Dempster-Shafer Theory of Evidence", Lecture Notes in Economics and Mathematical Systems No. 425., Springer-Verlag, 1995.

33. Kyburg H.E., Jr., "Bayesian and Non Bayesian Evidential Updating", Artificial Intelligence, Vol. 31, pp. 271-294, 1987.

34. Lowrance J.D., Garvey T.D.,"Evidential Reasoning: An Implementation for Multisensor Integration", technical Note 307, Artificial Intelligence Center, Computer Science and Technology Division, SRI International, Menlo Park, CA, 1983.

35. Lesh S.A.,"An Evidential Theory Approach to Judgment-Based Decision Making", Ph.D. Thesis, Dept. of Forestry and Environmental Studies, Duke University, Durham, NC, 1986.

36. Mahler R., "Combining Ambiguous Evidence with Respect to Abiguous a priori Knowledge, I: Boolean Logic", IEEE Trans. on SMC, Part 1: Systems and Humans, Vol. 26, No. 1, pp. 27-41, 1996.

37. Mesalkin L.D., "A Generalization of Sperner's Theorem on the Number of Subsets of a Finite Set", Theory Prob., 8, pp. 203-204, 1963.

38. Milner E.C., "A Combinatorial Theorem on Systems of Sets", J. London Math. Soc., 43, pp. 204-206, 1968.

39. Pearl J.,"Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", Computer Science, Artificial Intelligence, Morgan-Kauffman, 1988.

40. Revue pour la Science (french edition of Scientific American), "Einstein, le père du temps moderne", Issue no 11, May 2002.

41. Provan G.M.,"The validity of Dempster-Shafer Belief Functions", International Journal of Approximate Reasoning, Vol. 6, pp 389-399, 1992.

42. Rota G.C.,"Theory of Möbius Functions", Zeitung für Wahrscheinlichkeitstheorie und Verwandte Gebiete 2, pp. 340-368, 1964.

43. Ruspini E.H, "The Logical Foundation of Evidential Reasoning", Tech. Note 408, SRI International, Menlo park, CA, USA, 1986.

44. Ruspini E.H, Lowrance D.J., Start T.M.,"Understanding Evidential Reasoning", International Journal of Approximate Reasoning, Vol. 6, pp 401-424, 1992.

45. Schubert J.,"On Nonspecific Evidence", International Journal of Intelligent Systems, Vol. 8, pp. 711-725, 1993.

46. Semerdjiev E., Tchamova A.,"An Approach For a Fast reduction of the Full Ignorance in the Target Identification Process", Proceedings of First International Conference on Information Fusion (Fusion'98), Las Vegas, USA, July 1998.

47. Shafer G.,"A Mathematical Theory of Evidence", Princeton University Press, Princeton, New Jersey, 1976.

48. Shafer G.,"Constructive Decision Theory", Working paper, Dec. 1982 (basis for the following reference).

49. Shafer G.,Tversky A., "Languages and designs for probability", Cognitive Sc., Vol.9, pp. 309-339, 1985.

50. Shafer G., "The Combination of Evidence", Int. Journal of Intelligent Systems, Vol. 1, pp155-179, 1986.

51. Shafer G., Pearl J.,"Readings in Uncertain Reasoning", Morgan-Kauffman, 1990.

52. Shafer G.,"Perspectives on the Theory and Practice of Belief Functions", International Journal of Approximate Reasoning, Vol. 4, pp 323-362, 1990.

53. Shapiro, "On the Counting Problem for Monotone Boolean Functions", Comm. Pure Appl. Math., 23, pp. 299-312, 1970.

54. Smarandache F.,"A Unifying Field in Logics: Neutrosophic Logic. Neutrosophy, Neutrosophic Set, Probability, and Statistics", (Second Edition), American Research Press, Rehoboth, 2000.

55. Smarandache F. (Editor),"Proceedings of the First International Conference on Neutrosophics", Univ. of New Mexico, Gallup Campus, NM, USA, 1-3 December 2001, Xiquan, Phoenix, 2002.

56. Smets Ph.,"Constructing the pignistic probability function in a context of uncertainty", in Uncertainty in Artificial Intelligence 5, Henrion, M. and Shachter, R. D. and Kanal, L. N. and Lemmer, J.F. (Editors), North Holland, Amsterdam, pp. 29-40, 1990.

57. Smets Ph.,"The Combination of Evidence in the Transferable Belief Model", IEEE Trans. on PAMI, Vol. 12, no. 5, 1990.

58. Smets Ph.,"Quantifying Beliefs by Belief Functions: An Axiomatic Justification", Proc. of 13th International Joint Conf. on Artificial Intelligence, Chambéry, France, pp. 598-603, 1993.

59. Smets Ph.,"Belief Function: The Disjunctive Rule of Combination and the Generalized Bayesian Theorem", International Journal of Approximate Reasoning, Vol. 9, pp. 1-35, 1993.

60. Smets Ph.,"The alpha-junctions: combination operators applicable to belief function", Qualitative and quantitative practical reasoning, Springer, Gabbay D.M and Kruse R. and Nonnengart A. and Ohlbach H.J. Editors, pp. 131-153, 1997.

61. Smets Ph.,"The transferable belief model for quantified belief representation", Handbook of Defeasible Reasoning and Uncertainty Management Systems, D. M. Gabbay and Ph. Smets (Editors), Vol. 1, Kluwer, Doordrecht, The Netherlands, 1998.

62. Smets Ph.,"Data Fusion in the Transferable Belief Model", Proceedings of 3rd Int. Conf. on Inf. Fusion (Fusion 2000), pp. PS-21–PS33, (http://www.onera.fr/fusion2000), Paris, July 10-13, 2000

63. Sosnowski Z.A., Walijewski J.,"Generating Fuzzy Decision Rules with the use of Dempster-Shafer Theory", http://cksr.ac.bialystok.pl/jwal/papers/esm/fds-esm.html.

64. Sperner E.,"Ein Satz über Untermengen einer endlichen Menge", Math Z., 27, pp. 544-548, 1928.

65. Sudano J.,"Inverse pignistic Probability Transforms", in Proc. of the 5th Intern. Conf. on Inf. Fusion (Fusion 2002), Annapolis, MD, USA, pp. 763-768, July 8-11, 2002.

66. Sudano J.,"The System Probability Information Content (PIC) Relationship to Contributing Components, Combining Independent Multi-Source Belief, Hybrid and Pedigree Pignistic Probabilities", in Proc. of the 5th Intern. Conf. on Inf. Fusion (Fusion 2002), Annapolis, MD, USA, pp. 1277-1283, July 8-11, 2002.

67. Sun H., he K., Zhang B.,"The Performance of Fusion Judgment on Dempster-Shafer Rule", Chinese Journal of Electronics, Vol. 8, no. 1, Jan. 1999.

68. Tribus M.,"Rational, Descriptions, Decisions and Designs", Pergamon Press Inc., 1969 (French Transl.,"Decisions Rationnelles dans l'Incertain", Masson et Cie, Paris, 1972) .

69. Voorbraak F.,"On the Justification of Dempster's rule of Combination", Artificial Intelligence, Vol. 48, pp. 171-197, 1991.

70. Wierzchon S.T., Klopotek M.A.,"Evidential Reasoning. An Interpretative Investigation", Wydawnictwo Akademii Podlaskiej Publisher, PL ISSN 0860-2719, 304 pages, Poland, February 2002 (http://www.ipipan.waw.pl/ klopotek/mak/book2a.htm).

71. Yager R., "Entropy and Specificity in a Mathematical Theory of Evidence" Int. J. General Systems, Vol. 9, pp. 249-260, 1983.

72. Yager R., Kacprzyk J., Fedrizzi M. (Editors),"Advances in the Dempster-Shafer Theory of Evidence" ISBN: 0-471-55248-8, Hardcover 608 Pages, February 1994

73. Zadeh L.A.,"A Theory of Approximate Reasoning" Machine Intelligence, J. Hayes, D. Michie and L. Mikulich Eds, Vol. 9, pp. 149-194, 1979.

74. Zadeh L.A.,"On the Validity of Dempster's rule of Combination of Evidence" Memo M 79/24, Univ. of California, Berkeley, 1979.