An introduction to the 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 presents the basic mathematical settings of 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. Within this framework, 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 two 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. With the development of computer science, the last half of the 20th century has became very prolific for the development 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 *avant-gardiste* neutrosophy unifying theory by F. Smarandache (2000). This paper is a brief introduction of the new theory of plausible and paradoxical reasoning developped by the author which can be interpreted as a generalization of the theory of evidence. Due to space limitation, only a very short presentation of the Dempster-Shafer theory will be presented in the next section to help to set up the foundations of our new theory in section 3. The full presentation of this theory is presented in [5]. A discussion on the justification of the new rule of combination of uncertain and paradoxical sources of evidences will appear also in section 3. Two simple illustrative examples of the power and usefulness of this new theory will also presented at the end of this paper. The mathematical foundations of this new theory can be found in [5].

2 The Dempster-Shafer theory of evidence

The Dempster-Shafer theory of evidence (DST) is usually considered as a generalization of the bayesian theory of subjective probability [10] and offers a simple and direct representation of ignorance [15]. The DST has shown its compatibility with the classical probability theory, with boolean logic and has a feasible computational complexity for problems of small dimension. It is a powerful theoretical tool which can be applied for the representation of incomplete knowledge, belief updating, and for combination of evidence through the Demspter's rule of combination.

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 is called the frame of discernment of hypotheses or universe of discourse. 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 noted $\mathcal{P}(\Theta) = 2^\Theta$ 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*

$$m(\emptyset) = 0 \qquad \textit{and} \qquad \sum_{A \in 2^\Theta} m(A) \equiv \sum_{A \subseteq \Theta} m(A) = 1 \tag{1}$$

$m(A)$ corresponds to the measure of the partial 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 both be 0 even if $m(A_1 \cup A_2) \neq 0$. Even more peculiar, note that $A \subset B \not\Rightarrow m(A) < m(B)$ (i.e. $m(.)$ is not monotone to inclusion). Hence, the bba $m(.)$ is in general different from a probability distribution $p(.)$.

2.2 Belief and plausibility functions

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

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{2}$$

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

It can been shown [10] that a belief function Bel(.) can be characterized without reference to the information granule $m(.)$ and that from any given belief function Bel(.), one can always associate an unique information granule $m(.)$ from the Möbius inversion formula.

Definition 3. *The plausibility Pl(A) of any assertion $A \subset 2^{\Theta}$, which measures the total belief mass that can move into A (interpreted sometimes as the upper probability of A), is defined by*

$$Pl(A) \triangleq 1 - Bel(A^c) = \sum_{B \subseteq \Theta} m(B) - \sum_{B \subseteq A^c} m(B) = \sum_{B \cap A \neq \emptyset} m(B) \tag{3}$$

Bel(A) summarizes all our reasons to believe in A and Pl(A) expresses how much we could believe in A. Let now $(\Theta, m(.))$ be a source of information, then it is always possible to build the following *pignistic* probability [3, 16] (bayesian belief function) by choosing $\forall \theta_i \in \Theta$, $P\{\theta_i\} = \sum_{B \subseteq \Theta | \theta_i \in B} \frac{1}{|B|} m(B)$. One always gets

$$\forall A \subseteq \Theta, \qquad \text{Bel}(A) \quad \leq \quad [P(A) = \sum_{\theta_i \in A} P\{\theta_i\}] \quad \leq \quad \text{Pl}(A) \tag{4}$$

2.3 The Dempster's rule of combination

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

$$m(C) \triangleq [m_1 \oplus m_2](C) = \frac{\sum_{A \cap B = C} m_1(A) m_2(B)}{1 - \sum_{A \cap B = \emptyset} m_1(A) m_2(B)} \tag{5}$$

$m(.)$ is a proper bba if $K \triangleq= 1 - k \equiv 1 - \sum_{A \cap B = \emptyset} m_1(A) m_2(B) \neq 0$. The quantity k is called the *weight of conflict* between the bodies of evidences. When $K = 0$ (i.e. $k = 1$), $m(.)$ *does not exist* and the bodies of evidences $\mathcal{B}_1$ and $\mathcal{B}_2$ are said to be totally contradictory. Such case arises whenever the cores of Bel$_1$(.) and Bel$_2$(.) are disjoint. The same problem of existence has already been pointed out in the presentation of the optimal Bayesian fusion rule in [4].

The Dempster's rule of combination proposed by G. Shafer in [10] has been strongly criticized by the disparagers of the DST in the past decades because it had not been completly well justified by the author in his book, even if this has been corrected later in [11]. The DS rule is now accepted since the axiomatic of the transferable belief model (TBM) developed by Smets in [13, 7, 8, 14, 15] from an idea initiated by Cheng and Kashyap in [1]. The Dempster's and the optimal bayesian fusion rules [4] coincide exactly when $m_1(.)$ and $m_2(.)$ become bayesian basic probability assignments and if we accept the principle of indifference within the optimal Bayesian fusion rule. Many numerical examples of the Dempster's rule of combination can be found in [10]. What is more interesting now, is to focuse our attention on the following disturbing example.

Example 1. A simple but disturbing example

In 1982, Lofti Zadeh 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

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

The DS rule yields the unexpected result $m(T) = \frac{0.0001}{1-0.0099-0.0099-0.9801} = 1$ which means that the patient suffers with certainty from brain tumor !!!. This unexpected result arises from the fact that the two 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.

3 A new theory for plausible and paradoxical reasoning

3.1 Presentation

As seen in previous 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 Dempter's rule of combination. In most of practical applications based on the DST, some ad-hoc or heuristic recipes must always 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 conclusions (or cannot provide a reliable results at all). Even if nowadays, the DST provides fruithfull results in many applications, we strongly argue that this theory is still too limited because it is based on the two following restrictive constraints as already reported in literature

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 do not allow us to deal with the more general and practical problems involving uncertain reasoning and the fusion of uncertain, imprecise and paradoxical sources of information. The constraint $C1$ is very strong actually since it does not allow paradoxes between elements of the frame of discernment Θ. The DST accepts as foundation the commonly adopted principle of the third exclude. Even if at first glance, it makes sense in the traditional classical thought, we present here a new theory which does not accept this principle of the third exclude and accepts and deals with paradoxes.

The constraint C1 assumes that each elementary hypothesis of 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 Θ satisfying C1 even for some very simple problems where each elementary hypothesis corresponds to a vague concept or attributes. In such cases, the elementary elements of Θ cannot be precisely separated without ambiguity such that no refinement of Θ satisfying the first constraint is possible. Our second remark concerns the universal nature of the frame of discernment. It is clear that, in general, the *same* Θ is interpreted differently by the bodies of evidence or experts. Some subjectivity 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 corresponds to the C2 constraint. This vision seems to be too excessive because usually independent bodies of evidence provide their beliefs about some hypotheses only with respect to their own worlds of knowledge and experience 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 has to include the possibility to deal with evidences arising from different sources of information which don't have access to absolute interpretation of the elements Θ under consideration. This yields to accept the paradoxical information as the basis for a new general theory of probable reasoning. Actually, 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 as it will be seen. 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 deals with bba $m(.) \in [0,1]$ such that $m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) = 1$
- our general theory deals with new bba $m(.) \in [0,1]$ 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 exhaustivity assumption 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 and in the following, that Θ characterizes an exhaustive frame of discernment. Θ will be called a *general* frame of discernment in the sequel to emphaze the fact that Θ does not satisfy the Dempster-Shafer C1 constraint.

Definition 4. *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 $(A \cap B) \in D^\Theta$.*

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 [2]. 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 although this problem is more than one century old ...We just know that the cardinality numbers of D^Θ follow the Dedekind's numbers (minus one) when $\text{Card}(\Theta) = n$ increases, i.e. $\mid D^\Theta \mid = 1, 2, 5, 19, 167, 7580, 7828353, \ldots$ when $\text{Card}(\Theta) = n = 0, 1, 2, 3, 4, 5, 6, \ldots$ 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 unfortunately not possible in general as already argued.

Example 2.

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 $\Theta = \{\theta_1, \theta_2, \theta_3\}$,

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

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

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

Definition 5. *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*

$$m(\emptyset) = 0 \qquad and \qquad \sum_{A \in D^\Theta} m(A) = 1 \tag{6}$$

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

$$Bel(A) = \sum_{B \in D^\Theta, B \subseteq A} m(B) \qquad and \qquad Pl(A) = \sum_{B \in D^\Theta, B \cap A \neq \emptyset} m(B) \tag{7}$$

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 $\forall A \in D^\Theta, \mathrm{Bel}(A) \leq \mathrm{Pl}(A)$.

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

$$\forall i = 1, \ldots, n \qquad P\{\theta_i\} = \sum_{A \in D^\Theta} \alpha_{\theta_i}(A) m(A) \tag{8}$$

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 $\alpha_{\theta_i}(A) = \frac{1}{|A|}$ when $\theta_i \subseteq A$. 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 3. If $\Theta = \{\theta_1, \theta_2, \theta_3\}$ then $P\{\theta_1\}$ equals

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

Same kind of expressions can be derived for $P\{\theta_2\}$ and $P\{\theta_2\}$. 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 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.

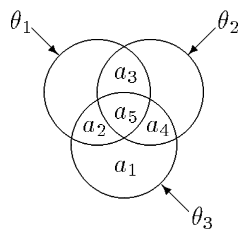

Fig.1 : Representation of $A = (\theta_1 \cap \theta_2) \cup \theta_3 \equiv a_1 \cup a_2 \cup a_3 \cup a_4 \cup a_5$

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 similar process.

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 7. *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*

$$\forall C \in D^{\Theta}, \qquad m(C) \triangleq [m_1 \oplus m_2](C) = \sum_{A,B \in D^{\Theta}, A \cap B = C} m_1(A) m_2(B) \tag{9}$$

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 statisfying (6). The global belief function Bel(.) is then obtained from the granule $m(.)$ through (7). This rule of combination is commutative and associative and can always be used for fusion of rational or paradoxical sources of information. 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. The theoretical justification of our rule of combination can be obtained as in [17] by the maximization of the joint entropy of the two paradoxical sources of information. This justification is reported in details in the companion paper [5]. The important result is that any fusion of sources of information generates either uncertainties, paradoxes or more generally both. This is intrinsic to the general fusion process itself. This general fusion rule can also be used within the 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 $\sum m(A) = 1$. In practice, for the sake of fair comparison between several alternatives or choices, it is better and simplier 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.

3.6 Zadeh's example revisited

Let's take back the disturbing Zadeh's example 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 (9), 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

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

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 ....

3.7 Mahler's example revisited

Let's consider now the following example excerpt from the R. Mahler's paper [9]. We consider that our classification knowledge base consists of the three (imaginary) new and rare diseases corresponding to following frame of discernment

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

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

$$\begin{array}{lll} m_0(\theta_1) = 0.2 & m_0(\theta_2) = 0.2 & m_0(\theta_3) = 0.2 \\ m_0(\theta_2 \cap \theta_3) = 0.2 & m_0(\theta_1 \cap \theta_2) = 0.1 & m_0(\theta_1 \cap \theta_2 \cap \theta_3) = 0.1 \end{array}$$

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 (9) 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

$$\begin{array}{lll} m(\theta_1) = 0.002 & m(\theta_2) = 0.200 & m(\theta_3) = 0.040 \\ m(\theta_1 \cap \theta_2) = 0.260 & m(\theta_2 \cap \theta_3) = 0.360 & m(\theta_1 \cap \theta_2 \cap \theta_3) = 0.100 \\ m(\theta_1 \cap (\theta_2 \cup \theta_3)) = 0.038 & & \end{array}$$

Therefore the final belief function given by (7) is

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

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.

4 Conclusion

In this paper, the foundations for a new theory of paradoxical and plausible reasoning have been shortly presented. This theory takes into account in the combination process itself the possibility for both 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 this general framework, no refinement is possible to apply directly the classical Dempster-Shafer theory of evidence. In our new theory, 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 shown that in general, the combination of evidence yields unavoidable paradoxes and, through two simple, but illustrative, examples, that conclusion drawn from this new theory, provides a result which agrees with the human reasoning and becomes useful to take a decision on complex problems where DST usually fails. A complete presentation of this new theory is available in [5].

References

1. 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.
2. Dedekind R.,"Über Zerlegungen von Zhalen durch ihre grössten gemeinsammen Teiler", In Gesammelte Werke, Bd. 1., pp. 103-148, 1897.
3. 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.
4. Dezert J., "Optimal Bayesian Fusion of Multiple Unreliable Classifiers", Proceedings of 4th Intern. Conf. on Information Fusion (Fusion 2001), Montréal, Aug. 7-10, 2001.
5. Dezert J., "Foundations for a new theory of plausible and paradoxical reasoning", To appear in next issue of the International Journal of Information & Security, edited by Prof. Tzv. Semerdjiev, CLPP, Bulgarian Academy of Sciences, 40 pages, 2002.
6. Dubois D., Prade H.,"Théories des Possibilités. Application à la Représentation des Connaissances en Informatique", Editions Masson, Paris, 1985.
7. Dubois D., Garbolino P., Kyburg H.E., Prade H., Smets, Ph.,"Quantified Uncertainty", J. Applied Non-Classical Logics, Vol.1, pp. 105-197, 1991.
8. 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.
9. 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.
10. Shafer G.,"A Mathematical Theory of Evidence", Princeton University Press, Princeton, New Jersey, 1976.
11. Shafer G.,Tversky A., "Languages and designs for probability", Cognitive Sc., Vol.9, pp. 309-339, 1985.
12. Smarandache F.,"An Unifying Field in Logics: Neutrosophic Logic", (Second Edition), American Research Press, Rehoboth, 2000 (ISBN 1-879585-76-6).
13. Smets Ph.,"The Combination of Evidence in the Transferable Belief Model", IEEE Trans. on PAMI, Vol. 12, no. 5, 1990.
14. 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.
15. 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.
16. 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
17. 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.
18. Zadeh L.A.,"A Theory of Approximate Reasoning" Machine Intelligence, J. Hayes, D. Michie and L. Mikulich Eds, Vol. 9, pp. 149-194, 1979.