The DSmT approach for information fusion and some open problems

Jean Dezert [a,1] and Florentin Smarandache [b]
[a] *ONERA, 92320 Châtillon, France*
[b] *Dept. of Math., Univ. of New Mexico, USA*

Abstract. This paper introduces the recent theory of plausible and paradoxical reasoning, known as DSmT (Dezert-Smarandache Theory) in the literature, which deals with imprecise, uncertain and potentially highly conflicting sources of information. Recent publications have shown the interest and the potential ability of DSmT to solve fusion problems where Dempster-Shafer Theory (DST) provides counter-intuitive results, especially when conflict between sources becomes high and information becomes vague and imprecise. This short paper presents the foundations of DSmT, its main rules of combination including the most recent ones and introduce briefly some open challenging problems in fusion.

Keywords. Information fusion, Dezert-Smarandache theory, DSmT, Plausible reasoning

1. Introduction

The development of the DSmT [8] arises from the necessity to overcome the inherent limitations of the DST [7] which are closely related with the acceptance of Shafer's model (i.e. working with an *homogeneous*[1] frame of discernment Θ defined as a finite set of *exhaustive* and *exclusive* hypotheses θ_i, $i = 1, \dots, n$), the third middle excluded principle, and Dempster's rule for the combination of independent sources of evidence. Limitations of DST are well reported in literature [17,13] and several alternative rules to Dempster's rule of combination can be found in [1,16,3,5,6,8]. DSmT provides a new mathematical framework for information fusion which appears less restrictive and more general than the basis and constraints of DST. The basis of DSmT is the refutation of the principle of the third excluded middle and Shafer's model in general, since for a wide class of fusion problems the hypotheses one has to deal with, can have different intrinsic nature and also appear only vague and imprecise in such a way that precise refinement is just impossible to obtain in reality so that the exclusive elements θ_i cannot be properly identified and defined. Many problems involving fuzzy/vague continuous and rela-

[1]Email addresses: jean.dezert@onera.fr, smarand@unm.edu.

[1]Although the homogeneity of Θ is not explicitly mentioned in the DST, it is a strong implicit assumption inherent to the Shafer's model. When working with DST, one implicitly assumes that all finite and exclusive elements of Θ have somehow the same semantic nature, otherwise the complement defined over the power-set becomes just a non-sense. The Shafer's model cannot deal directly with non-homogeneous elements (carrying different semantics) of Θ. This property however is necessary in many applications where the information given by the sources can't be expressed with same semantic due to the potentially different intrinsic nature of information carried by the sources/experts/sensors.

tive[2] concepts described in natural language with different semantic contents and having no absolute interpretation enter in this category. We claim that in general, the negation/complement is not accessible, but DSmT offers the possibility to deal with negation and Shafer's model as well. When the model of the problem fits with these constraints (negation follows from exclusivity constraints), we include them in the frame and then one forms the hyper-power set in the normal way. Thus DSmT deals naturally with negations/complements when necessary. DSmT starts with the notion of *free DSm model* and considers Θ only as a frame of exhaustive elements which can potentially overlap and have different intrinsic semantic natures and which also can change with time with new information and evidences received on the model itself. DSmT offers a flexibility on the structure of the model one has to deal with. When the free DSm model holds, the conjunctive consensus is performed. If the free model does not fit the reality because it is known that some subsets of Θ contain elements truly exclusive but also possibly truly non existing at all at a given time (in dynamic[3] fusion), new fusion rules must be performed to take into account these integrity constraints. The constraints can be explicitly introduced into the free DSm model to fit it adequately with our current knowledge of the reality; we actually construct a *hybrid DSm model* on which the combination will be efficiently performed. Shafer's model, which is the basis of DST, corresponds to a very specific hybrid DSm (and homogeneous) model including all possible exclusivity constraints. DSmT has been developed to work with any kind of model, to combine imprecise, uncertain and potentially high conflicting sources for static and dynamic information fusion. DSmT refutes the idea that sources provide their beliefs with the same absolute interpretation of elements of Θ; what is considered as good for somebody can be considered as bad for somebody else. Advances and first applications of DSmT are detailed in [8].

2. Notion of hyper-power set

Let $\Theta = \{\theta_1, \ldots, \theta_n\}$ be a finite set (called frame) of n exhaustive elements[4]. The free Dedekind's lattice denoted *hyper-power set* D^Θ [8] is defined as

1. $\emptyset, \theta_1, \ldots, \theta_n \in D^\Theta$.
2. If $A, B \in D^\Theta$, then $A \cap B$ and $A \cup B$ belong to D^Θ.
3. No other elements belong to D^Θ, except those obtained by using rules 1 or 2.

If $|\Theta| = n$, then $|D^\Theta| \leq 2^{2^n}$. The generation of D^Θ is presented in [8]. Since for any given finite set Θ, $|D^\Theta| \geq |2^\Theta|$, we call D^Θ the *hyper-power set* of Θ. $|D^\Theta|$ for $n \geq 1$ follows the sequence of Dedekind's numbers:1,2,5,19,167,7580,... An analytical expression of Dedekind's numbers obtained by Tombak and al. can be found in [8].

[2] The notion of relativity comes from the own interpretation of the elements of the frame Θ by each sources of evidences involved in the fusion process.

[3] i.e. when the frame Θ and/or the model $\mathcal{M}$ is changing with time.

[4] We do not assume here that elements θ_i have the same intrinsic nature and are necessary exclusive. There is no restriction on θ_i but the exhaustivity.

3. Free and hybrid DSm models

$\Theta = \{\theta_1, \ldots, \theta_n\}$ denotes the finite set of hypotheses/concepts characterizing the fusion problem. D^Θ constitutes the *free DSm model* $\mathcal{M}^f(\Theta)$ and allows to work with fuzzy concepts which depict a continuous and relative intrinsic nature. Such kinds of concepts cannot be precisely refined in an absolute interpretation because of the unapproachable universal truth. When all θ_i are truly exclusive discrete elements, D^Θ reduces naturally to the classical power set 2^Θ. This is what we call the Shafer's model. We deonote it $\mathcal{M}^0(\Theta)$. Between the free DSm model and the Shafer's model, there exists a wide class of fusion problems represented in term of DSm hybrid models where Θ involves both fuzzy continuous concepts and discrete hypotheses. In such class, some exclusivity constraints and possibly some non-existential constraints (especially when working on dynamic fusion) have to be taken into account. Each hybrid fusion problem is then characterized by a proper hybrid DSm model $\mathcal{M}(\Theta)$ with $\mathcal{M}(\Theta) \neq \mathcal{M}^f(\Theta)$ and $\mathcal{M}(\Theta) \neq \mathcal{M}^0(\Theta)$. From a general frame Θ, we define a map $m(.) : D^\Theta \rightarrow [0,1]$ associated to a given body of evidence $\mathcal{B}$ as

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

$m(A)$ is the *generalized basic belief assignment/mass* (gbba) of A. The *generalized belief and plausibility functions* are defined as:

$$\text{Bel}(A) \triangleq \sum_{\substack{B \subseteq A \\ B \in D^\Theta}} m(B) \quad \text{Pl}(A) \triangleq \sum_{\substack{B \cap A \neq \emptyset \\ B \in D^\Theta}} m(B) \tag{2}$$

4. Classic DSm fusion rule

When the free DSm model holds, the conjunctive consensus, called DSm classic rule (DSmC), is performed on D^Θ. DSmC of two independent[5] sources associated with gbba $m_1(.)$ and $m_2(.)$ is thus given $\forall C \in D^\Theta$ by [8]:

$$m_{\mathcal{M}^f(\Theta)}(C) \equiv m(C) = \sum_{\substack{A,B \in D^\Theta \\ A \cap B = C}} m_1(A) m_2(B) \tag{3}$$

Since D^Θ is closed under $\cup$ and $\cap$ set operators, DSmC guarantees that $m(.)$ is a proper generalized belief assignment, i.e. $m(.) : D^\Theta \rightarrow [0,1]$. DSmC is commutative and associative and can always be used for the fusion of sources involving fuzzy concepts whenever the free DSm model $\mathcal{M}^f(\Theta)$ holds. This rule can be directly and easily extended for the combination of $k > 2$ independent sources [8].

---

[5]While independence is a difficult concept to define in all theories managing epistemic uncertainty, we consider that two sources of evidence are independent (i.e. distinct and noninteracting) if each leaves one totally ignorant about the particular value the other will take.

5. Hybrid DSm fusion rule

When $\mathcal{M}^f(\Theta)$ does not hold (some integrity constraints exist), one deals with a proper DSm hybrid model $\mathcal{M}(\Theta) \neq \mathcal{M}^f(\Theta)$.The first general rule working on any model has been called DSm hybrid rule (DSmH) in [8]. More sophisticated rules based on differnt proportional conflict redistributions have recently been proposed [9] and only the most efficient one is presented in section 7. DSmH for $k \geq 2$ sources is defined for all $A \in D^\Theta$ as :

$$m_{\mathcal{M}(\Theta)}(A) \triangleq \phi(A) \cdot \Big[S_1(A) + S_2(A) + S_3(A)\Big] \tag{4}$$

where $\phi(A)$ is the *characteristic non-emptiness function* of a set A, i.e. $\phi(A) = 1$ if $A \notin \boldsymbol{\emptyset}$ and $\phi(A) = 0$ otherwise, where $\boldsymbol{\emptyset} \triangleq \{\boldsymbol{\emptyset}_{\mathcal{M}}, \emptyset\}$. $\boldsymbol{\emptyset}_{\mathcal{M}}$ is the set of all elements of D^Θ which have been forced to be empty through the constraints of the model $\mathcal{M}$ and $\emptyset$ is the classical/universal empty set. $S_1(A) \equiv m_{\mathcal{M}^f(\theta)}(A)$, $S_2(A)$, $S_3(A)$ are defined by

$$S_1(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ (X_1 \cap X_2 \cap \ldots \cap X_k) = A}} \prod_{i=1}^{k} m_i(X_i) \tag{5}$$

$$S_2(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in \boldsymbol{\emptyset} \\ [\mathcal{U}=A] \vee [(\mathcal{U} \in \boldsymbol{\emptyset}) \wedge (A = I_t)]}} \prod_{i=1}^{k} m_i(X_i) \tag{6}$$

$$S_3(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_k \in D^\Theta \\ u(c(X_1 \cap X_2 \cap \ldots \cap X_k)) = A \\ (X_1 \cap X_2 \cap \ldots \cap X_k) \in \boldsymbol{\emptyset}}} \prod_{i=1}^{k} m_i(X_i) \tag{7}$$

with $\mathcal{U} \triangleq u(X_1) \cup \ldots \cup u(X_k)$ where $u(X)$ is the union of all θ_i that compose X, $I_t \triangleq \theta_1 \cup \ldots \cup \theta_n$ is the total ignorance, and $c(X)$ is the conjunctive normal form[6] of X. $S_1(A)$ corresponds to DSmC rule for k independent sources based on $\mathcal{M}^f(\Theta)$; $S_2(A)$ represents the mass of all relatively and absolutely empty sets which is transferred to the total or relative ignorances associated with non existential constraints (if any, like in some dynamic problems); $S_3(A)$ transfers the sum of relatively empty sets directly onto the canonical disjunctive form of non-empty sets. DSmH generalizes DSmC and is not equivalent to Dempster's rule. It works for any models (the free DSm model, Shafer's model or any other hybrid models) when manipulating *precise* generalized (or eventually classical) basic belief functions.

[6]In Boolean algebra the conjunctive normal form is a conjunction of disjunctions, in its simplest form, which is unique; in this paper we consider each disjunction formed by a singleton or by a union of singletons; for example: $A \cap B \cap (C \cup D)$ is a conjunctive normal form; also, $X = (A \cup B) \cap C \cap (A \cup C)$ is a conjunction of disjunctions, but it is not in its simplest form, then its conjunctive normal form is $c(X) = (A \cup B) \cap C$ since $C \cap (A \cup C) = C$. The conjunctive normal form is introduced here in order to improve the original formula given in [8] for preserving the neutral impact of the vacuous belief mass $m(\Theta) = 1$ within complex hybrid models.

6. Fusion of imprecise beliefs

Since it difficult to have sources/human experts providing precise beliefs, a more flexible theory dealing with imprecise information is necessary. So we extended DSmT for dealing with *admissible imprecise generalized basic belief* $m^I(.)$ defined as real subunitary intervals of $[0,1]$, or even more general as real subunitary sets (not necessarily intervals). These sets can be unions of (closed, open, or half-open/half-closed) intervals and/or scalars all in $[0,1]$. An imprecise belief assignment $m^I(.)$ over D^Θ is said *admissible* if and only if there exists for every $X \in D^\Theta$ at least one real number $m(X) \in m^I(X)$ such that $\sum_{X\in D^\Theta} m(X) = 1$. The following simple operators on sets (addition $\boxplus$ and multiplication $\boxdot$) are necessary [8] for the fusion of imprecise beliefs:

$$\mathcal{X}_1 \boxplus \mathcal{X}_2 \triangleq \{x \mid x = x_1 + x_2, x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2\}$$

$$\mathcal{X}_1 \boxdot \mathcal{X}_2 \triangleq \{x \mid x = x_1 \cdot x_2, x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2\}$$

From these operators, one generalizes DSmC from scalars to sets as follows [8] (Chap. 6): $\forall A \neq \emptyset \in D^\Theta$,

$$m^I_{\mathcal{M}^f(\Theta)}(A) = \boxed{\sum_{\substack{X_1,X_2,\ldots,X_k\in D^\Theta \\ (X_1\cap X_2\cap\ldots\cap X_k)=A}}} \boxed{\prod_{i=1,\ldots,k}} m_i^I(X_i) \tag{8}$$

where $\boxed{\sum}$ and $\boxed{\prod}$ represent the summation, and respectively product, of sets. The DSmH fusion of imprecise beliefs takes a form similar to (4), except that $m_{\mathcal{M}(\Theta)}(A)$, $S_1(A)$, $S_2(A)$ and $S_3(A)$ have to be replaced by $m^I_{\mathcal{M}(\Theta)}(A)$, $S_1^I(A)$, $S_2^I(A)$ and $S_3^I(A)$ respectively, and where also the classical product $\cdot$ and sum $+$ operators have to be replaced by their corresponding operators on sets, i.e. $\boxdot$ and $\boxplus$. The definitions of $S_1^I(A)$, $S_2^I(A)$ and $S_3^I(A)$ are similar to (5)-(7) except that $\sum$ and $\prod$ operators are replaced by $\boxed{\sum}$ and $\boxed{\prod}$ and $m_i(X_i)$ by $m_i^I(X_i)$. A detailed presentation on DSmH imprecise with several examples can be found in [8].

7. New proportional conflict redistribution rule

DSmH is one of possible issues for the fusion of highly conflicting vague imprecise and uncertain information, but DSmH is not the unique solution for such fusion[7] and more complex and efficient rules can also be used instead. Due to space limitations, we present here the most sophisticated Proportional Conflict Redistribution (PCR) rule we developed so far. Let's $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ be the frame of the problem, G denoting the hyper-power set D^Θ or classical power set 2^Θ depending on the model one wants to deal with, and two belief assignments $m_1, m_2 : G \to [0,1]$ such that $\sum_{X\in G} m_i(X) = 1$, $i = 1, 2$. The general principle of the PCR rules is to compute

[7] while DSmH provides a satisfactory mathematical solution for any model and for static or dynamic fusion problems.

1. the conjunctive rule, $\forall X \in G$, $m_{1\ldots s}(X) = \sum_{\substack{X_1,\ldots,X_s \in G \\ X_1 \cap \ldots \cap X_s = X}} \prod_{i=1}^{s} m_i(X_i)$
2. the conflicting masses (partial and/or total), The *total conflicting mass* k_{12} drawn from two sources is defined by $k_{12} = \sum m_1(X_1)m_2(X_2) = \sum m(X_1 \cap X_2)$ where the sum is for all $X_1, X_2 \in G$ such that $X_1 \cap X_2 = \emptyset$. The total conflicting mass is nothing but the sum of *partial conflicting masses* $m(X_1 \cap X_2)$, where $X_1 \cap X_2 = \emptyset$, represents a partial conflict between X_1 and X_2. These formulas can be generalized for $s \geq 2$ sources [9].
3. the proportional redistribution of the conflicting mass (total or partial) to non-empty sets involved in the model according to all integrity constraints.

The way the conflicting mass is redistributed yields actually to five versions of PCR rules denoted PCR1, PCR2, ...PCR5 [9]. PCR rules work for any degree of conflict $k_{12} \in [0,1]$, for any model. They work both in DST and DSmT and for static or dynamical fusion and can be directly extended for the fusion of imprecise belief as well. The sophistication/complexity (but correctness) of proportional conflict redistribution increases from PCR1up to the PCR5 presented here. For static fusion, PCR1 coincides with the Weighted Average Operator [4], but PCR1 and WAO do not preserve the neutral impact of the vacuous belief assignment (VBA) $m_v(\theta_1 \cap \ldots \cap \theta_n) = 1$ in the fusion, i.e. when $s > 1$, $[m_1 \oplus \ldots \oplus m_s \oplus m_v](X)$ becomes different from $[m_1 \oplus \ldots \oplus m_s](X)$ although $m_v(.)$ brings no extra specific information. PCR2-PRC5 overcome this drawback and preserve the neutral impact of VBA. A detailed presentation of PCR rules with many examples can be found in [9]. PCR5 fusion rule for two sources is is given by [**?**]: $\forall X \in G \setminus \{\emptyset\}$

$$m_{PCR5}(X) = m_{12}(X) + \sum_{\substack{Y \in G \setminus \{X\} \\ c(X \cap Y) = \emptyset}} [\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)}]$$

where $c(x)$ is the conjunctive normal form of x, $m_{12}(.)$ is the conjunctive consensus operator and where all denominators are *different from zero*. If a denominator is zero, that fraction is discarded. The general PCR5 formula for $s \geq 2$ sources is given in [9]. PCR5 rule is quite natural since it redistributes proportionally the partial conflicting mass only to the elements involved in the partial conflict, considering the conjunctive normal form of the partial conflict. PCR5 proposes in the authors opinions a more exact redistribution of conflicting mass to non-empty sets following the logic of the conjunctive rule. Furthermore, improvements of DSmH and PCR rules based on degrees of intersection, union and inclusion are also possible as already introduced in [11] but will not be reported here.

8. Open challenging problems

There are many open fundamental and theoretical problems related with reasoning under uncertainty. We briefly introduce here what we consider today as major open challenging problems for future. Most of these problems have already been attacked over the years by the research community but no clear solution and consensus have arisen so far in our opinion. So we deeply think that more research effort have definitly to be put on these important problems for improving the reasoning under uncertainty and for the development of the next generation of performant multisensor systems involving uncertain, incomplete, imprecise and conflicting information.

1. The first fundamental question concerns the characterization of any type of source in term of information content and consequently the development of an unified theory of uncertainty [18]. The only consensus available today seems to be the Shanon entropy for Bayesian sources. Some attempts to characterize other type of sources (fuzzy sources, evidential sources, parodoxist sources, etc) have already been proposed by example in [15,2,14,12], but no general theory exists today to quantify the measure of uncertainty for all kind of sources. We even don't know if one or several measures are necessary for such purpose. Many different measures have been proposed and are in competition actually.
2. The second challenging problem concerns the fusion of qualitative and symbolic information and the fusion of qualitative information with quantitative information in expert systems.
3. The third major open problem is the development of a general theory of decision from reasoning under uncertainty. Several decision theories (mainly based on probability theory) have been developed, but a general unified decision theory is still missing.

9. Conclusion

This paper brings a short overview on DSmT and the last advances obtained in the development of new fusion rules. The major rules of fusion have been presented here for the combination of precise and imprecise beliefs and open challenging and very difficult general problems have been proposed for future research and for probably several generations of researchers in this field.

References

[1] Dubois D., Prade H., *On the unicity of Dempster rule of combination*, International Journal of Intelligent Systems, Vol. 1, pp 133-142, 1986.

[2] Harmanec D., *Measure of uncertainty and information*, Imprecise Probability Project,1999 (http://ippserv.rug.ac.be/home/ipp.html)

[3] Inagaki T., *Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory*, IEEE Trans. on reliability, Vol. 40, no. 2, pp. 182-188, 1991.

[4] Jøsang, A., Daniel, M., Vannoorenberghe, P., *Strategies for Combining Conflicting Dogmatic Beliefs*, Proceedings of the 6th International Conference on International Fusion, Vol. 2, pp. 1133-1140, 2003.

[5] Lefèvre E., Colot O., Vannoorenberghe P., *Belief functions combination and conflict management*, Inf. Fusion J., Vol. 3, No. 2, pp. 149-162, 2002.

[6] Sentz K., Ferson S., *Combination of evidence in Dempster-Shafer Theory*, SANDIA Tech. Report, SAND2002-0835, 96 pages, April 2002.

[7] Shafer G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[8] Smarandache F., Dezert J. (Editors), *Applications and Advances of DSmT for Information Fusion*, Am. Res. Press, Rehoboth, 2004, http://www.gallup.unm.edu/~smarandache/DSmT-book1.pdf.

[9] Smarandache F., Dezert J., *Proportional Conflict Redistribution Rules for Information Fusion*, submitted to JAIF Journal, http://arxiv.org/pdf/cs.AI/0408064, March 2005.

[10] Smarandache F., Dezert J., *Information Fusion Based on New Proportional Conflict Redistribution Rules*, to appear in Proc. of Fusion 2005 Int. Conf., Philadelphia, July 2005.

[11] Smarandache F., Dezert J., *The combination of paradoxical, uncertain and imprecise sources of information based on DSmT and neutro-fuzzy inference*, Journal of Applied Mathematics & Computing, Seoul, South Korea, March 2005, http://xxx.lanl.gov/PS_cache/cs/pdf/0412/0412091.pdf.

[12] Liu C., Grenier D., Jousselme A.-L., Bossé E., *Measuring ambiguity in the evidence theory*, Submitted to IEEE Trans on SMC, Part A , 2003.

[13] Voorbraak F., *On the justification of Dempster's rule of combination*, Artificial Intelligence, 48, pp. 171-197, 1991.

[14] Walley P., *Measures of uncertainty in expert systems*, Art. Intel., 83, pp. 1-58, 1996.

[15] Yager R.R., *Entropy and specificity in a mathematical theory of evidence*, Int. J. Generaal Systems, Vol. 9, pp. 249–260, 1983.

[16] Yager R.R., *On the Dempster-Shafer framework and new combination rules*, Information Sciences, Vol. 41, pp. 93–138, 1987.

[17] Zadeh L., *On the validity of Dempster's rule of combination*, Memo M 79/24, Univ. of California, Berkeley, 1979.

[18] Zadeh L., *Toward a unified theory of uncertainty*, IPMU conference, 2004.