Reply to the Comments of R. Haenni on the paper "Belief functions combination and conflict management"

E. Lefevre[1], O. Colot[2] and P. Vannoorenberghe[3]

[1] *PSI, EA 2120, Université & INSA de Rouen*
Place Emile Blondel, F-76131 Mont Saint Aignan Cedex
Eric.Lefevre@insa-rouen.fr
[2] *IRCOM-SIC, UMR CNRS 6615, Bat. SP2MI*
BP 30179, Bd Marie et Pierre Curie, F-86962 Chasseneuil du Poitou Cedex
colot@sic.sp2mi.univ-poitiers.fr
[3] *HEUDIASYC, UMR 6599 CNRS, Université de Technologie de Compiègne*
BP 20529, F-60205 Compiègne Cedex, France
Patrick.Vannoorenberghe@hds.utc.fr, http://www.hds.utc.fr/~pvannoo

---

The paper untitled "Belief functions combination and conflict management" has been published in a recent issue of the journal [1]. The problem of information combination is introduced in the context of Dempster-Shafer theory of evidence or belief function theory [2]. The authors adopt the Transferable Belief Model (TBM) point of view, a non-probabilistic interpretation of the theory. This model has been introduced by P. Smets [3] and its use for information fusion has been demonstrated in a lot of applications. We would like to express our thanks to R. Haenni for his comments [4] and to the Editor for the opportunity to reply to these comments. We first recall the problem of uncertain data combination and the method proposed in [1] to aggregate belief functions. Solutions to normalize the belief function after combination are detailed in section 2. In section 3, main differences between TBM and the Probabilistic Argumentation Systems (PAS), proposed by R. Haenni in [5], are presented. Finally, we give our opinion concerning the challenges of Dempster-Shafer theory in the section 4. Practical applications of the TBM for pattern recognition problems are referenced.

1 Problem

The problem is to combine belief functions $\mathbf{m}$ defined on a same frame of discernment Θ in order to obtain a single belief function m which summarizes the available information. In [1], a generic formulation is proposed:

$$m(A) = m_{\cap}(A) + w(A, \mathbf{m}).m(\emptyset) \qquad \forall\, A \subseteq \Theta \tag{1}$$

where $w(A, \mathbf{m})$ are coefficients, called weighting factors, which must satisfy:

$$\sum_{A \subseteq \Theta} w(A, \mathbf{m}) = 1. \tag{2}$$

The idea consists in retaining the conjunctive part of the evidence $m_{\cap}$ and to eventually add a weighted part of the conflict $m(\emptyset)$. In these equations, the number of weighting factors w is $2^{|\Theta|}$ coefficients. This number does not depend of the number of belief functions to be combined. In [1], the authors present several ways to adjust these weighting factors. It has been proved that adjusting judiciously these coefficients allows rewriting existing operators involved in the literature (including Dempster's rule, Yager's rule, ...). Unfortunately, with this formulation (equation (1)) and as mentioned by R. Haenni, the operator loss the associativity property but has other interesting advantages.

2 Why or how normalize ?

In this section, we summarize the main solutions to manage the conflict existing after the combination of several belief functions. Why or how normalize? As far as we know, three solutions can be envisaged.

As mentioned by G. Shafer in its original works on belief functions [2], the first solution which consists to distribute the conflict proportionally among the subsets of the frame of discernment seems to be reasonable. It allows keeping intact the properties of the belief functions. This solution can be justified in the Probabilistic Argumentation Systems but not in the Transferable Belief Model.

Another solution has been investigated by P. Smets [6]. It consists to keep the conflict as it is and, by this way, to consider that the frame of discernment must be extended by an additional element that covers all other possible states of the world. Thus, $\emptyset$ can be interpreted as one or several hypotheses which are not taken into account in the initial frame, this is the so-called "Open World" assumption [1]. A similar solution has been proposed by R. Yager in [7] which postulates on the introduction of a new element in the frame of discernment.

Concerning the third solution, let us consider m_1 and m_2 two belief functions and $m_1 \oplus m_2$ their combination with Dempster's rule. If the result is

[1] In the TBM, the condition $\sum_{\emptyset \neq A \subseteq \Theta} = 1$ is not assumed and $m(\emptyset) > 0$ is accepted.

counter-intuitive, of course we can refine the model in introducing discounting coefficients α_1 and α_2. And what about the values of these coefficients? In addition, these two coefficients are respectively attached to each belief function. The point of view defended in [1] is quite different because weighting factors w involved in the combination operator have to be optimized for each subset of the frame of discernment. With these two approaches, additional information is needed to adjust the parameters (α in the first solution or w in the second). In [1], the relationships between these two approaches (discounting and weighted combination) has been explicitly proved (see Appendix). In fact, weighting factors w can be obtained for each subset of the frame from discounting coefficients α and belief functions $\mathbf{m}$ involved in the combination. In addition, in a recent paper [8], the authors demonstrate that they give different results in a pattern recognition problem and show that adjusting weighting factors w improves significantly the percentage of correct classification, *Quod erat demonstrandum.*

3 Comparing the TBM and the PAS

Can we consider that the TBM and the PAS model are similar models? Our answer is no as they don't share the same 'semantic' [2] .

In the TBM, there is a domain Y on which the agent, denoted You, has some beliefs or weighted opinions about which of the elements of Y is the actual one. You can bet on Y, by what we mean that You can bet on which of the elements of Y is the actual one. These quantified beliefs are represented by a belief function, and their operational definition is derived from Your betting behavior.

In the PAS model, there is a domain X (a set of assumptions) on which there is a probability measure, and a one-to-many mapping from the X domain to a Y domain (a set of hypotheses). Given the special nature of the mapping, one can only express upper and lower limits about what might be the probabilities about the value of the actual world in Y. It happens that the lower envelop of these probabilities is a belief function, and it is assumed that this function represents Your quantified beliefs about which element of Y is the actual one.

Mathematically, there is a link between the TBM and the PAS model as, for any belief function defined within the TBM, one can build a X space endowed with a probability measure, and thus deduce the same structure as considered in the PAS model. But this is just a mathematical similarity. It is the same as the one that links hydraulic flux and electric flux. Both share the same

[2] Ph. Smets, Personal communication.

equations, still when you need water, you turn the faucet on, not the light.

The difference appears when one ask for the meaning of the 'probability' on X in the PAS model? In subjective probability theory, its meaning is to be found in a betting behavior. So in the PAS model, one must be ready to bet not only on the actual value of Y (as in the TBM), but also on the actual value of X. This is where TBM and the PAS model diverge. In the TBM, the concept of betting on the actual value of X is meaningless, in which case the 'probability' on X is also meaningless.

Thus the discrepancy between the two models comes from the fact that in the PAS model one can bet on both the actual value of X and the actual value of Y, whereas in the TBM, one can only bet on the actual value of Y, the X space being just an artificial construct without any link to reality.

If for any reason, there is nevertheless a meaningful space X (and the one-to-many mapping from X to Y), then in the TBM, the beliefs on X would be represented by a belief function, not by a probability function.

4 Challenges

The final comment concerns future works on Dempster-Shafer theory. Of course, practical applications of such approaches (including all interpretations of belief functions) to uncertainty management are a real challenge. In this context, important efforts have been undertaken to induce belief functions from statistical data. For pattern recognition problems, classification rules based on the Dempster-Shafer theory of evidence have been introduced by several authors [9,10]. Three methods can be distinguished including the consonant method initially introduced by Shafer in the more general context of statistical inference [2], Appriou's separable method [10] and the distance-based classifier introduced by Denœux [11]. A comparison of such evidential classifiers has been made in [12]. Other approaches investigate the induction of decision trees based on credal inference [13,14]. They allow handling uncertain labels in pattern recognition problems and then constituting practical applications of the Dempster-Shafer theory to machine learning.

References

[1] E. Lefevre, O. Colot, P. Vannoorenberghe, Belief functions combination and conflict management, Information Fusion 3 (2) (2002) 149–162.

[2] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[3] P. Smets, What is Dempster-Shafer's model?, in: R. Yager, M. Fedrizzi, J. Kacprzyk (Eds.), Advances in the Dempster-Shafer Theory of Evidence, Wiley, 1994, pp. 5–34.

[4] R. Haenni, Are alternatives to Dempster's rule of combination real alternatives? comments on "About the belief function combination and the conflict management problem", Information Fusion 3 (3) (2002) 237–239.

[5] R. Haenni, N. Lehmann, Probabilistic argumentation systems: A new perspective on Dempster-Shafer theory, International Journal of Intelligent Systems (Special Issue on Dempster-Shafer theory of evidence) To appear (2002) xx–xx.

[6] P. Smets, The combination of evidence in the transferable belief model, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (5) (1990) 447–458.

[7] R. Yager, Hedging in the combination of evidence, Journal of Information and Optimization Sciences 4 (1) (1983) 73–81.

[8] E. Lefevre, O. Colot, P. Vannoorenberghe, D. Brucq, Informations et combinaison : Les liaisons conflictuelles, Traitement du Signal 18 (3) (2002) 161–177.

[9] L. Zouhal, T. Denœux, An evidence-theoretic k-nn rule with parameter optimization, IEEE Transactions on Systems, Man and Cybernetics-Part C 28 (2) (1998) 263–271.

[10] S. Fabre, A. Appriou, X. Briottet, Presentation and description of two classification methods using data fusion based on sensor management, Information Fusion 2 (2001) 49–71.

[11] T. Denœux, A neural network classifier based on Dempster-Shafer theory, IEEE Transactions on Systems, Man and Cybernetics, Part A : Systems and humans 30 (2) (2000) 131–150.

[12] P. Vannoorenberghe, T. Denœux, Likelihood-based vs. distance-based evidential classifiers, in: Proceedings of FUZZ-IEEE'2001, IEEE, Melbourne, Australia, 2001.

[13] P. Vannoorenberghe, T. Denœux, Handling uncertain labels in multiclass problems using belief decision trees, in: Proceedings of IPMU'2002, Annecy, France, 2002, pp. 1919–1926.

[14] Z. Elouedi, K. Mellouli, P. Smets, Belief decision trees: Theoretical foundations, International Journal of Approximate Reasoning 28 (2001) 91–124.