Max-Ent in fast belief fusion

M. Pappalardo and F. Villecco

Abstract. This paper deals with Jaynes' MaxEnt Principle and Yager's combination rule for weighting the masses associated with the focal elements. The main goal is to manage uncertainty by suitably modelling the degrees of belief. The solution of this problem is classically approached by the Bayesian method based on probability functions. The DS theory also provides a useful framework for the representation of information content in an uncertain variable. However in both approaches the weights, associated with the focal elements can be referred to probabilities, and must be know in advance. We show that, under Max-Entropy principle, the fusion of belief is faster. We present a new way of obtaining from two independent and equally reliable sources of evidence, a new set with the probability allotted on each focal element.

M.S.C. 2000: 03B52, 03B53, 03B42, 03B45.
Key words: belief, MaxEnt, fuzzy set, fast belief fusion.

1 Introduction

Let be $P(A)$ the probability of A, the Bayes' theory requires the relation $P(A) + P(\overline{A}) = 1$ in which between lack of disbelief and disbelief there is no distinction. Often in engineering design it is proper to do the fundamental change of replacing the precise value that a probability has with the concept that a probability has a degree of variability in an interval that provides a lower and upper bound. The idea of upper and lower probability, in belief functions, was proposed for handling uncertainty connected with subjectivity. The belief functions are a bridge between various models handling different forms of uncertainty. When there is not enough information on which to evaluate a probability, or when the information is non-specific, ambiguous, or in conflict, then the Bayesian model cannot be used. A method for handling data in presence of uncertainty with qualitative values is the theory of Dempster-Shafer (DS). The DS model includes the Bayesian probability as special case, and introduces the belief function as lower probabilities and the plausibility function as upper probabilities. The numerical measure, in presence of uncertainty, may be assigned to set of

BSG Proceedings 15. The International Conference "Differential Geometry - Dynamical Systems" DGDS-2007, October 5-7, 2007, Bucharest-Romania, pp. 154-162.

elements as well as to a single element. In the DS model the probabilities, apportioned to subsets and the mass v_i can move over each element. Let be the frame of discernment the next finite non-empty set $\Theta = \{x_1, ..x_n\}$, Θ is the set of all hypothesis. The basic probability is assigned in the range$[0, 1]$ to the 2^nsubset of Θ consisting of a singleton or conjunction of singleton of n basic elements x_i. The basic probability, a function which assigns the weight to the subset of the frame of discernment, is the mass function $m\,(.)$. The mass $m\,(\theta)$ s where we assign the probability that we are unable to assign otherwise. If the belief remain apportioned in single elements, then the DS model corresponds to the Bayesian model of probability. Formally the description of basic probability assignments can be represented with the next equations:

$$\begin{cases} m : P\,(X) \rightarrow [0,1] \ ; \ (\,X : universal\,set) \\ \sum_{A\subseteq\Theta} m\,(A_i) = 1 \\ m\left(\varnothing\right) = 0 \ ; \ \left(\varnothing \ : \ \textit{empty set}\right) \end{cases}$$

If $m_1$and$m_2$are basic probabilities from the independent evidence, and$\{A_{1\mathrm{i}}\}$, $\{A_{2\mathrm{j}}\}$ the sets of focal points, then Dempster's model of combination gives the rule of fusion. Given two basic probabilities from the independent evidence, if it is

$$\sum\nolimits_{A_{1i}\cap A_{2j}\neq\varnothing} m_1(A_i)m_2(A_j) > 0 \quad ; \quad A_k \neq \varnothing$$

then following Dempster's rule

$$m\,(A) = (m_1 \oplus m_2)\,(A) \stackrel{def}{=} \frac{\sum_{A_{1i}\cap A_{2j}=A_k} m_1(A_i)m_2(A_j)}{1 - \sum_{A_{1i}\cap A_{2j}=\varnothing} m_1(A_i)m_2(A_j)}$$

combines two or more probabilities.

The Dempster's rule is easy to use and gives a quick mathematical model for handling uncertainty including Bayesian theory. The reliability of result depends on the interpretation of the basic probability assignment. When the conflict K,

$$K = \sum\nolimits_{A_{1i}\cap A_{2j}=\varnothing} m_1\,(A_i)m_2\,(A_j)$$

between the sources of independent basic probability, becomes important then the DS rules present some limitations.

The DS rules present some weakness, more than once reported by Zadeh, because if the conflict $K$is important, then the result of fusion is unacceptable. The rules are mainly based on the extension of the domain of the probability functions. In the applications there exists many cases where DS rules assign low belief to elements of sets with larger cardinality [1].

2 Zadeh's example

Zadeh example was stated in 1986 [2, 4] and describes the conflict in the Dempster Shafer method which leads to a paradox. Shortly the example is as follows [3]. Two doctors examine An ill person, during a health check up, is examined by two Physicians . Each doctor make a diagnosis that the patient suffers from either meningitis

(M), bruise (C) or brain tumor (T). Thus the set of beliefs is Θ= {M,C, T }. Let us assume that both doctors agree in their low expectation of a tumor, but disagree in likely cause and provide the following diagnosis

$$\text{m1(M)} = 0.99 \text{ m1(T)} = 0.01 \text{ and m2(C)} = 0.99 \text{ m2(T)} = 0.01.$$

According to Dempster's rule, by combining the two basic belief functions, one gets the unexpected final conclusion

$$m\left(t\right) = \frac{0.0001}{1 - 0.0099 - 0.0099 - 0.9801} = 1$$

which means that the patient suffers with a full certainty from brain tumor !!!. This paradox comes from the fact that the two physicians (bodies of evidence) agree that the patient most probably does not suffer from tumor pain but are in almost full contradiction for the other causes of the disease. This example shows the many constraints of the practical use of the DST for automatic reasoning.

Several authors tried to explain the anomaly of the result of Dempster's rule of combination in this example. Due to the high degree of conflict which is present in a such extreme case, and pointed out by Zadeh in order to show the weakness of this rule, it is often argued that in such case the result of Dempster's rule must not be taken directly without checking the level of the conflict between sources of evidence. This is trivially true but there not exist yet a theoretical explanation to decide beforehand if one can trust or not the result of such rule of combination, especially in complex systems involving many sources and many hypotheses.

Many algorithms have been suggested, and many alternative rules have been proposed to overcome the difficulties of the computational complexity of reasoning and to avoid the limitations of Dempster's rule. A greater number of algorithms can be represented by changing the fusion's rules and by choosing the solutions according to the application and the need of capturing the epistemic uncertainty. Many rules are justified or criticized (even rejected) but all of them surely show that there exists a great number of possible rules of combinations. The computation of upper$P^*(A_j)$ and lower$P_*(A_j)$ probabilities has the same dual interpretation as in the standard Bayesian calculus. If we don't accept that the DS rules assign certainty to element of sets with lower cardinality, than we have to reject the rules of the probabilistic functions.

The basic probability $m = 2^\Theta \to [0,1]$ assign a numerical value to the focal elements$m\left(A_i\right)$. If we reduce the focal elements of the frame of discernment $\Theta = \{x_1, ..x_n\}$ into singletons basic elements, then the fusion takes the same structure of Bayesian rule.

In a fusion, if n_A is the power of set $\Theta_A$and n_B is the power of set Θ_B, then the power of the set resulting from a fusion is

$$\Theta_{A\cap B} = \Theta_A \oplus \ \Theta_B \to n_{A\cap B} = Min\left(n_A, n_B\right).$$

Many limitations of fusion rule are due to the evaluation of the independence of the two distributions$m_1\left(A_{1i}\right)$ and $m_2\left(A_{2j}\right)$.

The problem of the independence is a crucial point in combining evidence. Given the events $\theta = \{A, B, C\}$, the elements $2^{\Theta} = 2^3$ of the frame of discernment of all hypotheses are :

$$\left\{A, B, C, A \cup B, A \cup C, B \cup C, A \cup B \cup C, \varnothing\right\} \quad .$$

The masses of probability of a distribution are:

$$m(.) = (m_1 \oplus m_2)(A) = \{m(A), m(B), m(C), m(A \cup B), \\ m(A \cup C), m(B \cup C), m(A \cup B \cup C)\}$$

The masses $\{m(A), m(B), m(C)\}$ are uncoupled values of probability whereas the masses$\{m(A \cup B), m(A \cup C), m(B \cup C), m(A \cup B \cup C)\}$ are coupled values of distribution. In the fusion system, the reduction of uncertainty and complexity is the central problem because the resulting data are the input design parameter for many applications, especially the real time. We need to assign the probability to a singleton basic elements with uncoupled distribution. The goal of this paper is the definition of a rule for decoupling distributions on the base of the MaxEnt Principle.

3 Yager's rule

Yager's rule of combination: Yager admits that in case of conflict the result is not reliable, so that $k12$ plays the role of an absolute discounting term added to the weight of ignorance.

The commutative (but not associative) Yager rule, denoted here by index Y is given by:

$$\begin{cases} m_Y\left(\varnothing\right) = 0 \\ m_Y(A) = \sum\limits_{\substack{X, Y \in 2^{\Theta} \\ X \cap Y = A}} m_1(X)\, m_2(Y), \quad \forall A \in 2^{\Theta},\ A \neq \varnothing\ ,\ A \neq \Theta \\ m_Y(\Theta) = m_1(\Theta)\, m_2(\Theta) + \sum\limits_{\substack{X, Y \in 2^{\Theta} \\ X \cap Y = \varnothing}} m_1(X)\, m_2(Y), \quad \textit{when } A = \Theta \end{cases}$$

4 Examples

Let us consider the following distribution:

$$m_1(.) = \{m_1(A) = 0.2; m_1(B) = 0.3\}; m(C) = 0;\ m(A \cup B) = 0,1; m(B \cup C) = 0,4$$

We apply the D&S rule and MaxEnt.

By decoupling the set $m_1(.)$ we have:

		m1(ABC)	m1(AB)	m1(AC)	m1(BC)	m1(A)	m1(B)	m1(C)
		0	0.1	0	0.4	0.2	0.3	0
m(ABC)ME	0	0	0	0	0	0	0	0
m(AB)ME	0	0	0	0	0	0	0	0
m(AC)ME	0	0	0	0	0	0	0	0
m(BC)ME	0	0	0	0	0	0	0	0
m(A)ME	0.33333333	0	0.03333333	0	0.13333333	0.06666667	0.1	0
m(B)ME	0.33333333	0	0.03333333	0	0.13333333	0.06666667	0.1	0
m(C)ME	0.33333333	0	0.03333333	0	0.13333333	0.06666667	0.1	0

Figure 1: Distribution

$$m^{BME}(.) = \left[m_1(.) \oplus m^{ME}(.)\right] = \begin{bmatrix} A & B & C & A\cup B & B\cup C \\ 0.2 & 0.3 & 0 & 0.1 & 0.4 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \end{bmatrix} = \\ = \begin{bmatrix} A & B & C \\ 0.2 & 0.53 & 0.27 \end{bmatrix}$$

D.S.-MAXENT Model

m1(ABC)ME	m1(AB)ME	m1(AC)ME	m1(BC)ME	m1(A)ME	m1(B)ME	m1(C)ME
0	0	0	0	0.2	0.53333333	0.26666667

Figure 2: Solution

Now we apply Yager's rules with the same distribution:

$$m_1(.) = \{m_1(A) = 0.2; m_1(B) = 0.3; m(C) = 0; \\ m(A\cup B) = 0,1; m(B\cup C) = 0,4\}$$

$$m_y^{BME}(.) = \left[m_1(.) \oplus m^{ME}(.)\right] = \begin{bmatrix} A & B & C & A\cup B & B\cup C \\ 0.2 & 0.3 & 0 & 0.1 & 0.4 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \end{bmatrix} = \\ = \begin{bmatrix} A & B & C & A\cup B\cup C \\ 0.1 & 0.27 & 0.13 & 0.5 \end{bmatrix}$$

In the next figure we can see the two different results of D& S and Yager's rules

Let us now consider the Zadeh example, and apply the D& S and Yager's rules.

$$m_1(.) = \{m_1(A) = 0.9; m_1(B) = 0; m_1(C) = 0.1\};$$

$$m_2(.) = \{m_2(A) = 0; m_2(B) = 0.9; m_2(C) = 0.1\};$$

Yager-MAXENT Model						
$m_{y1}(ABC)_{ME}$	$m_{y1}(AB)_{ME}$	$m_{y1}(AC)_{ME}$	$m_{y1}(BC)_{ME}$	$m_{y1}(A)_{ME}$	$m_{y1}(B)_{ME}$	$m_{y1}(C)_{ME}$
0,500000	0	0	0	0,1	0,26666667	0,13333333

Figure 3: Solution

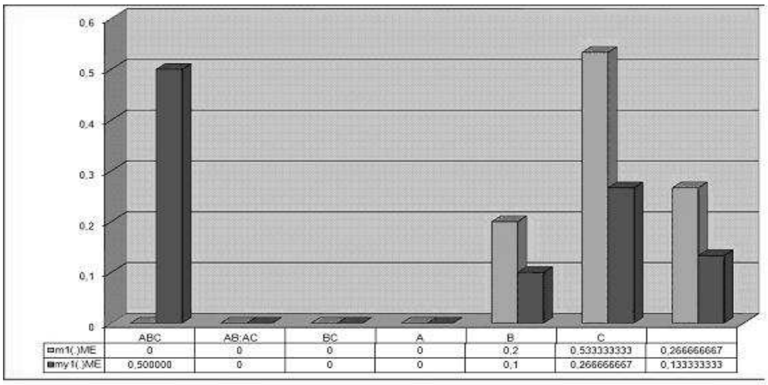

Figure 4: D&S and the Yager solutions

		$m_1(ABC)$	$m_1(AB)$	$m_1(AC)$	$m_1(BC)$	$m_1(A)$	$m_1(B)$	$m_1(C)$
		0	0	0	0	0,9	0	0,1
$m_2(ABC)$	0	0	0	0	0	0	0	0
$m_2(AB)$	0	0	0	0	0	0	0	0
$m_2(AC)$	0	0	0	0	0	0	0	0
$m_2(BC)$	0	0	0	0	0	0	0	0
$m_2(A)$	0	0	0	0	0	0	0	0
$m_2(B)$	0,9	0	0	0	0	0,81	0	0,09
$m_2(C)$	0,1	0	0	0	0	0,09	0	0,01

Figure 5: Distribution

D&S solution:

$$m_{DS}(.) = [m_1(.) \oplus m_2(.)] = \begin{bmatrix} A & B & C \\ 0.9 & 0 & 0.1 \\ 0 & 0.9 & 0.1 \end{bmatrix} = \begin{bmatrix} A & B & C \\ 0 & 0 & 1 \end{bmatrix}$$

		D.S. Model				
m12(ABC)	m12(AB)	m12(AC)	m12(BC)	m12(A)	m12(B)	m12(C)
0	0	0	0	0	0	1

Figure 6: Solution

Yager solution:

$$m_Y(.) = [m_1(.) \oplus m_2(.)] = \begin{bmatrix} A & B & C \\ 0.9 & 0 & 0.1 \\ 0 & 0.9 & 0.1 \end{bmatrix} = \begin{bmatrix} ABC & A & B & C \\ 0.99 & 0 & 0 & 0.01 \end{bmatrix}$$

		Yager Model				
my(ABC)	my(AB)	my(AC)	my(BC)	my(A)	my(B)	my(C)
0.990000	0	0	0	0	0	0.01

Figure 7: Solution

Applying Belief - MaxEnt to the Yager and D&S solution, we obtain: D&S

$$m^{BME}(.) = \left[m_{DS}(.) \oplus m^{ME}(.)\right] = \begin{bmatrix} A & B & C \\ 0 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} = \begin{bmatrix} A & B & C \\ 0 & 0 & 1 \end{bmatrix}$$

YAGER

$$\begin{aligned} m^{BME}(.) \quad &= \left[m_Y(.) \oplus m^{ME}(.)\right] = \begin{bmatrix} ABC & A & B & C \\ 0.99 & 0 & 0 & 0.1 \\ 0 & 1/3 & 1/3 & 1/3 \end{bmatrix} \\ &= \begin{bmatrix} ABC & A & B & C \\ 0.006667 & 0.33 & 0.33 & 0.333333 \end{bmatrix} \end{aligned}$$

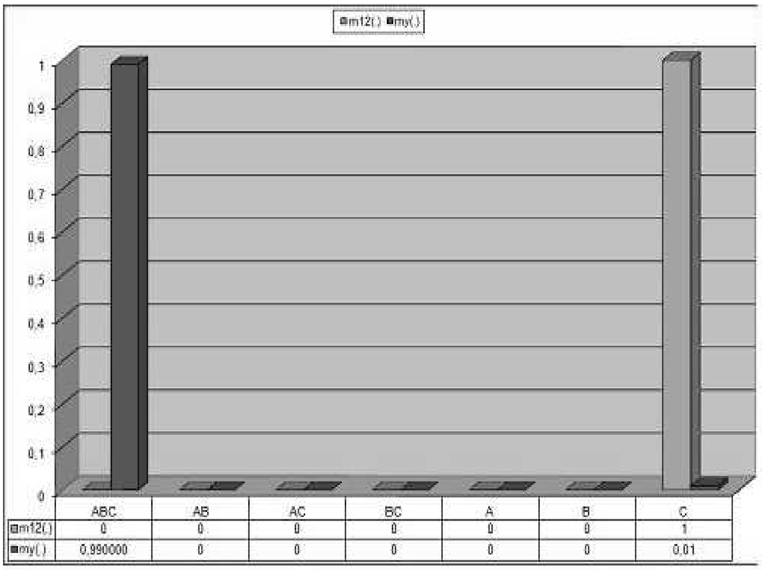

Figure 8: D&S and Yager's solution

D.S.-MAXENT Model						
m12(ABC)ME	m12(AB)ME	m12(AC)ME	m12(BC)ME	m12(A)ME	m12(B)ME	m12(C)ME
0	0	0	0	0	0	1

Figure 9: Solution

Yager-MAXENT Model						
my(ABC)ME	my(AB)ME	my(AC)ME	my(BC)ME	my(A)ME	my(B)ME	my(C)ME
0,006667	0	0	0	0,33	0,33	0,333333333

Figure 10: Solution

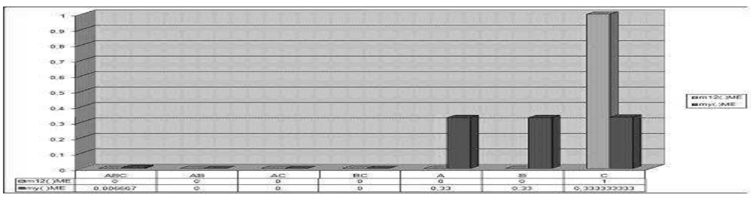

Figure 11: D&S and the Yager solutions

5 Conclusion

From the above results, based on the Belief-Maxent principle and D& S rules, we have shown that probability mass allots only on the basic elements , but it's impossible to overcome Zadeh's paradox.

By using, instead Yager's rule together with the Belief- Maxent principle it's impossible to decouple the distribution of masses, but we can overcome Zadeh's paradox, with a value of K, which measures the largeness of conflicting mass, equal to $K = 0,006667$.

References

[1] M. Pappalardo, *Maxent principle for handling uncertainty with qualitative values*, MAXENT 2006.

[2] F. Smarandache, J. Dezert, *Advances and Application of DSmt for Information Fusion*, American Research Press, 2004.

[3] L. Zadeh, *A simple view of the Dempster-Shafer theory of evidence and its implications for the rule of combination*, Berkeley Cognitive Science Report 33, 1985, University of California, Berkeley, CA; AI Magazine 7, 2 (1986), 85-90.

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

[5] L. Zadeh, *Review of Mathematical theory of evidence*, by Glenn Shafer, AI Magazine 5, 3 (1984), 81-83.

Authors' address:

Michele Pappalardo and Francesco Villecco
Mechanical Engineering, University of Salerno,
Via Ponte Don Melillo, 84084 Fisciano (SA), Italy.
E-mail: mpappalardo@unisa.it, fvillecco@unisa.it