

## On the properties of relative plausibilities

Fabio Cuzzolin

► **To cite this version:**

Fabio Cuzzolin. On the properties of relative plausibilities. IEEE Systems, Man, and Cybernetics Society (SMC'05), Oct 2005, Hawaii, United States. pp.594-599, 10.1109/ICSMC.2005.1571211 . inria-00590190

**HAL Id: inria-00590190**

**<https://hal.inria.fr/inria-00590190>**

Submitted on 6 May 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the properties of relative plausibilities

Fabio Cuzzolin

Department of Computer Science  
University of California at Los Angeles  
Los Angeles, CA, USA  
cuzzolin@cs.ucla.edu

**Abstract** - In this paper we investigate the properties of the relative plausibility function, the probability built by normalizing the plausibilities of singletons associated with a belief function. On one side, we stress how this probability is a perfect representative of the original belief function when combined with any arbitrary probability through Dempster's rule. This leads to conjecture that this function should also be the solution of the probabilistic approximation problem, formulated naturally in terms of Dempster's rule. On the other side, the geometric properties of relative plausibilities are studied in the context of the geometric approach to the theory of evidence, yielding a description of the representation property which suggests a sketch for the general proof of our conjecture.

**Keywords:** Belief functions, belief space, relative plausibilities, probabilistic approximation, Dempster's rule.

## 1 Introduction

The problem of finding the correct probabilistic approximation of belief functions has been widely studied in the last years. A number of papers have been published on this issue [13,14,15] (see [1] for a review), mainly in order to find efficient implementations of the rule of combination aiming to reduce the number of focal elements in coarsened frames [14] or using hierarchical clustering [15]. Tessem [11], for instance, incorporated only the highest-valued focal elements in his  $m_{ktx}$  approximation; a similar approach inspired the *summarization* technique formulated by Lowrance *et al.* [8]. The connection between belief functions and probabilities is as well the basement of a popular approach to the theory of evidence, Smets' *pignistic model* [10]. On his side, F. Voorbraak [12] proposed as solution the *relative plausibility* function  $\tilde{P}_s^*$ , i.e. the unique probability that, given a belief function  $s$  with plausibility  $P_s^*$ , assigns to each singleton its normalized plausibility. Afterwards, Cobb and Shenoy [16] described some properties of the relative plausibility of singletons and discussed its nature of probability function that is equivalent to the original belief function.

In this paper we will actually study some properties of this Bayesian function. Considering the central role of Dempster's rule of combination in the theory of evidence, we will formulate an approximation criterion based on

similarity between combinations. We will show how the relative plausibility function solves a “reduced” version of the problem, as a consequence of  $\tilde{P}_s^*$  being a “perfect” representative of  $s$  in the probabilistic region  $P$ , meaning that  $s \oplus t = \tilde{P}_s^* \oplus t$  for each probability  $t$  where  $\oplus$  denotes their Dempster's sum.

In the search for a general solution of this Dempster-based approximation problem, we will study the geometry of the relative plausibility function in the context of the geometric interpretation of the theory of evidence. The geometric behavior of the representation property, in particular, will suggest us a possible sketch for the general proof. The geometric analysis of belief and plausibility functions is due to the author. Black has also dedicated its doctoral thesis to the study of belief functions [2]. Another close reference could be a recent paper of Ha and Haddawy [7] in which they exploit methods of convex geometry to represent probability intervals.

## 2 The theory of evidence

### 2.1 Belief and plausibility functions

In the theory of evidence [6] [9] a basic probability assignment (b.p.a.) over a finite set  $\Theta$  (called *frame of discernment*) is a function  $m:2^\Theta \rightarrow [0,1]$  such that

$$m(\emptyset) = 0, \quad \sum_{A \subseteq \Theta} m(A) = 1, \quad m(A) \geq 0 \forall A \subseteq \Theta \quad (1)$$

Subsets of  $\Theta$  associated with non-zero values of  $m$  are called *focal elements* and their union *core*. The *belief function*  $s:2^\Theta \rightarrow [0,1]$  associated with the basic probability assignment  $m$  is defined as:

$$s(A) = \sum_{B \subseteq A} m(B). \quad (2)$$

Conversely, the basic probability assignment  $m$  of a belief function  $s$  can be uniquely recovered by means of the *Moebius inversion formula* [9] so that there is a 1–1 correspondence between the two set functions  $m \leftrightarrow s$ . In particular, a probability function is a peculiar belief

function which satisfies the additivity rule for disjoint sets (*Bayesian* b.f.).

Belief functions representing distinct bodies of evidence can be combined by means of *Dempster's rule* [6]. The *orthogonal sum* of two belief functions  $s_1, s_2$  is a new belief function  $s_1 \oplus s_2$  whose focal elements are all the possible intersections between pairs of focal elements of  $s_1$  and  $s_2$  respectively, and whose b.p.a. is

$$m(C) = \frac{\sum_{A \cap B = C} m_1(A) \cdot m_2(B)}{\sum_{A \cap B \neq \emptyset} m_1(A) \cdot m_2(B)}. \quad (3)$$

When all the intersections between focal elements of the two functions are empty, the denominator of Equation (3) goes to zero and we say that  $s_1$  and  $s_2$  are non-combinable. The definition easily extends to the combination of several belief functions. A dual representation of the evidence encoded by a belief function  $s$  is called *plausibility function*. Its value expresses the amount of evidence *not against* a proposition

$$P_s^*(A) = 1 - s(A^c) = \sum_{B \cap A \neq \emptyset} m(B) \quad (4)$$

where  $A^c$  is the complement of the set  $A$  in  $\Theta$ . From  $P_s^*$  we can derive a Bayesian belief function called *relative plausibility of singletons* by simply normalizing the plausibility values of all the elements of  $\Theta$ :

$$\begin{aligned} \tilde{P}_s^* \quad \Theta &\rightarrow [0,1] \\ x &\mapsto \tilde{P}_s^*(x) = \frac{P_s^*(x)}{\sum_{y \in \Theta} P_s^*(y)}. \end{aligned} \quad (5)$$

## 2.2 Belief and plausibility space

Motivated by the approximation problem, we recently introduced the language of convex geometry in the theory of evidence to study the geometry of belief functions [3,4,5]. Consider a frame of discernment  $\Theta$  and introduce in the Euclidean space  $\mathfrak{R}^N$ ,  $N = |\Theta| - 1$ , an orthonormal reference frame  $\{x_A\}$ ,  $A \subset \Theta$ ,  $A \neq \emptyset$ , in which each coordinate function  $x_A$  measures the belief value  $s(A)$  of  $A$ .

**Definition 1.** The *belief space* associated with  $\Theta$  is the set of points  $S_\Theta$  of  $\mathfrak{R}^N$  corresponding to a belief function.

We assume the domain  $\Theta$  fixed, and use the notation  $S$  to refer to the belief space. To determine which points of  $\mathfrak{R}^N$  "are" belief functions we can exploit the Moebius inversion lemma, by computing the corresponding b.p.a. and

checking the axioms  $m$  must obey. It is not difficult to prove [3,5] that  $S$  is convex. More precisely, if call

$$P_A = s \in S : m(A) = 1, \quad m(B) = 0 \forall B \neq A \quad (6)$$

the unique belief function assigning all the mass to a single subset  $A$  of  $\Theta$ , it can be proved that the belief space  $S$  coincides with the convex closure  $Cl$  of all the basis belief functions  $P_A$ , (see Figure 1)

$$S = Cl(P_A, A \subset \Theta, A \neq \emptyset). \quad (7)$$

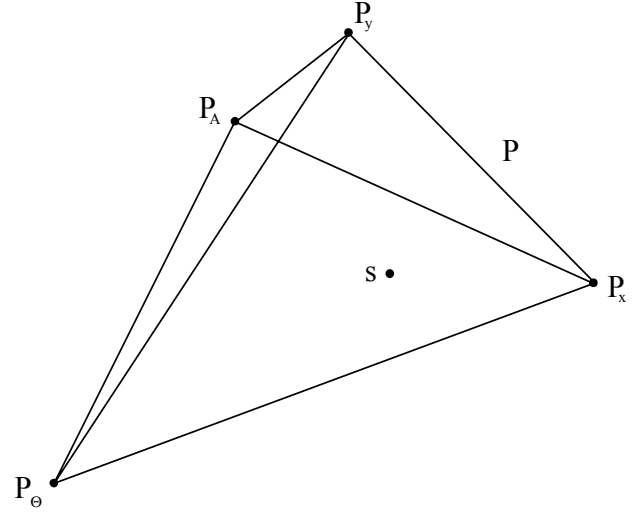


Figure 1. The belief space is a simplex whose vertices are the basis belief functions  $P_A$ .

Furthermore, any belief function  $s \in S$  can be written as a convex sum as follows:

$$s = \sum_{A \subset \Theta} m(A) P_A. \quad (8)$$

Since a probability is a belief function assigning non zero masses to singletons only, the set  $P$  of all the Bayesian belief functions is a subset of the boundary of  $S$ , precisely the simplex determined by all the basis functions associated with singletons:

$$P = Cl(P_x, x \in \Theta). \quad (9)$$

Analogously, we can call *plausibility space* the region  $\Pi$  of  $\mathfrak{R}^N$  whose points correspond to admissible plausibility functions. It can be proved [4] that  $\Pi$  is also a simplex,

$$\Pi = Cl(\Pi_A, A \subset \Theta) \quad (10)$$

whose vertices  $\Pi_A$  are exactly the plausibility functions associated with the basis belief functions  $P_A$ .

### 3 Dempster-based approximation

The rule of combination is central in the theory of evidence: belief functions are meaningful when combined with other b.f. in an inference process. We believe this should be taken into account when approaching the approximation problem. This led us to formulate a constraint on the "external" behavior of the desired approximation. *A good approximation, when combined with any other belief function, must produce results similar to what obtained by combining the original function.* Analytically, this can be expressed as

$$\hat{s} = \arg \min_{s \in C} \int_{t \in S} \text{dist}(s \oplus t, s' \oplus t) dt \quad (11)$$

where  $t \in S$  is an arbitrary belief function on the same frame, "dist" is a distance function in the Euclidean space (being the belief space a subset of  $\mathfrak{R}^N$ ), and  $C$  is the class of belief functions the approximation belongs to. Here we consider the class  $P$  of the Bayesian belief functions, and show how the relative plausibility function possess a peculiar property which candidates it to be the solution of the probabilistic approximation problem as posed in (11).

#### 3.1 Representation

Let us compute the general form of the difference  $s \oplus t - p \oplus t$ , where both  $p$  and  $t$  are Bayesian belief functions. To this extent we exploit the general *commutativity* property we have recently proved in [3]. We just saw that any belief function  $s$  can be expressed as a convex sum of simple b.f. (Equation (8)). As Dempster's rule is commutative with respect to convex combinations (see [3], Proposition 3), if  $t$  is a belief function with b.p.a.  $m_t$

$$s \oplus t = \sum_{A \subset \Theta} \frac{m_t(A) P_s^*(A)}{\sum_{B \subset \Theta} m_t(B) P_s^*(B)} \cdot s \oplus P_A \quad (12)$$

Immediately then, if  $t$  in particular is a Bayesian b.f.

$$s \oplus t = \sum_{x \in \Theta} \frac{t(x) P_s^*(x)}{\sum_{y \in \Theta} t(y) P_s^*(y)} \cdot P_x \quad (13)$$

$$p \oplus t = \sum_{x \in \Theta} \frac{t(x) p(x)}{\sum_{y \in \Theta} t(y) p(y)} \cdot P_x \quad (14)$$

where  $P_x$  is the basis belief function for the event  $\{x\}$ , since  $P_p^*(x) = p(x)$  as  $p$  is a probability. The difference  $s \oplus t - p \oplus t$  can hence be expressed as

$$\begin{aligned} & \sum_{x \in \Theta} t(x) P_x \cdot \left[ \frac{P_s^*(x)}{\Delta_{ts}} - \frac{p(x)}{\Delta_{tp}} \right] = \\ & = \frac{1}{\Delta_{ts} \Delta_{tp}} \sum_{x \in \Theta} t(x) P_x \cdot [P_s^*(x) \Delta_{tp} - p(x) \Delta_{ts}] \end{aligned} \quad (15)$$

having defined

$$\Delta_{ts} = \sum_{x \in \Theta} t(x) P_s^*(x), \quad \Delta_{tp} = \sum_{x \in \Theta} t(x) p(x). \quad (16)$$

**Theorem 1.** The relative plausibility of singletons  $\tilde{P}_s^*$  is a perfect representation of  $s$  in the probability subspace through Dempster's rule, i.e.  $\forall t \in P$

$$s \oplus t = \tilde{P}_s^* \oplus t. \quad (17)$$

*Proof.* Namely, if  $p(x) = \tilde{P}_s^*(x) = P_s^*(x) / \sum_{y \in \Theta} P_s^*(y)$

$$\begin{aligned} \Delta_{tp} &= \sum_{x \in \Theta} t(x) \frac{P_s^*(x)}{\sum_y P_s^*(y)} = \\ &= \frac{1}{\sum_y P_s^*(y)} \sum_{x \in \Theta} t(x) P_s^*(x) = \frac{\Delta_{ts}}{\sum_y P_s^*(y)} \end{aligned} \quad (18)$$

so that  $s \oplus t - \tilde{P}_s^* \oplus t$  becomes

$$\frac{\sum_x P_s^*(x)}{\Delta_{ts}^2} \sum_x t(x) P_x \cdot \left[ \frac{P_s^*(x) \Delta_{ts}}{\sum_y P_s^*(y)} - \frac{P_s^*(x) \Delta_{ts}}{\sum_y P_s^*(y)} \right] = 0. \quad (19)$$

#### 3.2 Reduced approximation criterion

An immediate consequence of Theorem 1 is that the modified version of the approximation problem in which the original b.f. is combined with all and only the Bayesian belief functions is trivially solved by  $\tilde{P}_s^*$ :

$$\tilde{P}_s^* = \arg \min_{p \in P} \int_{t \in P} \|s \oplus t - p \oplus t\| dt \quad (20)$$

whatever the chosen norm, being  $s \oplus t = \tilde{P}_s^* \oplus t$ . It is then natural to conjecture that the relative plausibility function could be a solution of the general approximation problem, too. An analysis of the geometry of  $\tilde{P}_s^*$  can possibly help us in this effort.



$$\begin{aligned}
\cdot [|X|-1] &= \sum_{X \cap \{x_1, \dots, x_k\} \neq \emptyset} m(X) [|X| - |X \cap \{x_1, \dots, x_k\}|] + \\
&+ \sum_{X \cap \{x_1, \dots, x_k\} = \emptyset} m(X) [|X| - 1] \geq 0
\end{aligned} \tag{28}$$

Also, for any chain of subsets  $\{x_1\} \subset \dots \subset \{x_1, \dots, x_k\} \subset \dots \subset \Theta$  the sequence  $\{\Sigma_k\}$  is decreasing, as

$$\begin{aligned}
\Sigma_{k+1} - \Sigma_k &= - \sum_{x_{k+1} \in X} m(X) + \sum_{X \cap \{x_1, \dots, x_{k+1}\} \neq \emptyset} m(X) + \\
&- \sum_{X \cap \{x_1, \dots, x_k\} \neq \emptyset} m(X) = - \sum_{X \supset x_{k+1}} m(X) + \\
&+ \sum_{X \supset x_{k+1}, \{x_1, \dots, x_k\} \subset X} m(X) = - \sum_{X \supset x_{k+1}, \exists i \in [1, \dots, k]: X \supset x_i} m(X) \leq 0.
\end{aligned} \tag{29}$$

In the binary case  $\tilde{P}_s^* = P_s^*$  and the relative plausibility function lays on the line  $(P_\Theta, P_s^*)$  as in Figure 2.

## 5 Towards a general solution

### 5.1 The geometry of representation

Figure 2 makes clear that  $(P_\Theta, P_s^*)$  corresponds to a *dual* line  $(\Pi_\Theta, s)$ , joining the belief function  $s$  with the vacuous plausibility  $\Pi_\Theta$ . The results of Section 4 can now be given a geometric interpretation in terms of this line, possibly enlightening our search for a solution of the general approximation problem (11).

**Theorem 4.** All the points  $t$  of the line  $(s, \tilde{P}_s^*)$  are perfect representatives of  $s$  when combined with arbitrary Bayesian functions,  $t \oplus p = s \oplus p \forall p \in P$ .

*Proof.* We can again exploit the results on Dempster's rule of affine combinations. If  $\sum_i \alpha_i = 1$ , [3]

$$s \oplus \sum_i \alpha_i s_i = \sum_i \beta_i s \oplus s_i \tag{30}$$

where  $\beta_i = \alpha_i \Delta_i / \sum_j \alpha_j \Delta_j$  and  $\Delta_i$  is the denominator of the b.p.a. of  $s \oplus s_i$  as in Section 3.1. Now, each point  $t$  on the line  $(s, \tilde{P}_s^*)$  can be written as

$$t = \lambda s + (1 - \lambda) \tilde{P}_s^* \tag{31}$$

for some  $\lambda \in [0, 1]$ . But then its combination with  $p$  is

$$\mu p \oplus s + (1 - \mu) p \oplus \tilde{P}_s^* \tag{32}$$

where the new line coordinate  $\mu$  is, according to (30),

$$\mu = \frac{\lambda \Delta_{sp}}{\lambda \Delta_{sp} + (1 - \lambda) \Delta_{\tilde{P}_s^* p}} \tag{33}$$

where the denominators of  $p \oplus s$  and  $p \oplus \tilde{P}_s^*$  are

$$\begin{aligned}
\Delta_{sp} &= \sum_x p(x) \sum_{A \supset x} m(A) = \sum_x p(x) P_s^*(x) \\
\Delta_{\tilde{P}_s^* p} &= \sum_x p(x) \tilde{P}_s^*(x) = \frac{\Delta_{sp}}{\sum_x P_s^*(x)}
\end{aligned} \tag{34}$$

so that, since  $s \oplus p = \tilde{P}_s^* \oplus p$  for Theorem 1,  $p \oplus t = p \oplus [\lambda s + (1 - \lambda) \tilde{P}_s^*] = \mu s \oplus p + (1 - \mu) s \oplus p = s \oplus p$ . In other words, the representatives of  $s$  in  $P$  are located on the line passing through  $s$  and  $\tilde{P}_s^*$ .

### 5.2 Probabilistic approximation in $S_2$

Let us see how the general criterion (11) works for the binary frame. The representation property allows us to rewrite the integral (11) in terms of the lines  $(s', \tilde{P}_s^*)$  (see Figure 3). Namely, after introducing the notation  $t = \tilde{P}_s^*$ ,

$$\begin{aligned}
&\int_{t \in P} \int_{s' \in (c, t)} |s \oplus s' - p \oplus t| ds' dt = \\
&= \int_{t \in P} \int_{\lambda \in [0, 1]} |\beta s \oplus t + (1 - \beta) s \oplus c - p \oplus t| d\lambda dt
\end{aligned} \tag{35}$$

since  $s \oplus s' = s \oplus [\lambda t + (1 - \lambda)c] = \beta s \oplus t + (1 - \beta) s \oplus c$  where  $\beta = \lambda \Delta_{st} / (\lambda \Delta_{st} + (1 - \lambda) \Delta_{sc})$  for (30) again. After observing that  $s \oplus (1, 1) = (1, 1)$  (a trivial consequence of its definition), and remembering that Dempster's rule maps a line passing through a pair of points into another line through the *images* of these points [3] we get that the images  $s \oplus c$  and  $s \oplus t$  also lay on a line passing through  $(1, 1)$ ! In other words,  $s \oplus t = \tilde{P}_{s \oplus c}^*$ . Now, if  $p = \tilde{P}_s^*$  expression (35) becomes

$$\int_{t \in P} \int_{\lambda \in [0, 1]} |1 - \beta| d\lambda |s \oplus c - \tilde{P}_s^* \oplus t| dt \tag{36}$$

since again  $s \oplus t = \tilde{P}_s^* \oplus t$ . Geometrically, all the difference vectors  $s \oplus s' - \tilde{P}_s^* \oplus t$  are *parallel* to the line  $(s \oplus c, s \oplus t)$ . This property is characteristic of the relative plausibility function, since for any other probability  $p$ ,  $p \oplus t$  does not lay in the intersection of this line with  $P$ . We can reasonably consider this elegant symmetry as an evidence of the correctness of our conjecture.

