Generalizations of the relative belief transform

Fabio Cuzzolin

Abstract Probability transformation of belief functions can be classified into different families, according to the operator they commute with. In particular, as they commute with Dempster's rule, relative plausibility and belief transforms form one such "epistemic" family, and possess natural rationales within Shafer's formulation of the theory of evidence. However, the relative belief transform only exists when some mass is assigned to singletons. We show here that relative belief is only a member of a class of "relative mass" mappings, which can be interpreted as low-cost proxies for both plausibility and pignistic transforms.

1 Introduction

The theory of evidence (ToE) [14] extends classical probability theory through the notion of *belief function* (b.f.), a mathematical entity which independently assigns probability values to *sets* of possibilities rather than single events. A belief function $b : 2^\Theta \rightarrow [0,1]$ on a finite set or *frame* Θ has the form $b(A) = \sum_{B \subseteq A} m_b(B)$, where the function $m_b : 2^\Theta \rightarrow [0,1]$ (called *basic probability assignment* or *basic belief assignment* b.b.a.) is both non-negative $m_b(A) \geq 0 \; \forall A \subseteq \Theta$ and normalized $\sum_{A \subseteq \Theta} m_b(A) = 1$. Subsets $A \subset \Theta$ associated with non-zero basic probabilities $m_b(A) \neq 0$ are called *focal elements*. A basic probability assignment m_b can be uniquely recovered from a belief function b by Moebius transform: $m_b(A) = \sum_{B \subseteq A} (-1)^{|A-B|} b(B)$. Special belief functions assigning non-zero masses to singletons only ($m_b(A) = 0$ whenever $|A| > 1$, $A \subseteq \Theta$) are called *Bayesian* b.f.s, and are in 1-1 correspondence with probability distributions on Θ. Different operators have been proposed for the combination of two or more belief functions, starting from the orthogonal sum originally formulated by A. Dempster [10].

Fabio Cuzzolin
Oxford Brookes University, Oxford, UK, e-mail: fabio.cuzzolin@brookes.ac.uk

Probability transformation of belief functions. The relation between belief and probability, in particular, has been an important subject of study in the theory of evidence, and a number of papers have been published on the issue of probability transform [9]. Many of these proposals, such as [13] or [18], seek efficient implementations of the rule of combination. A different, decision based approach to probability transformation is the foundation of Smets' "Transferable Belief Model" [15], in which decisions are made via the *pignistic probability* $BetP[b](x) = \sum_{A \supseteq \{x\}} \frac{m_b(A)}{|A|}$, justified via a number of rationality principles. Other proposals have been recently brought forward by Dezert et al. [11], Burger [1], Sudano [17] and others, based on redistribution processes similar to that of the pignistic transform. Two new Bayesian approximations of belief functions have been derived from purely geometric considerations [4] in the context of the geometric approach to the ToE [5], in which belief and probability measures are represented as points of a Cartesian space.

Relative plausibility and belief transforms. Following the efficient implementation approach, Voorbraak [19] has developed a probabilistic approximation intended to limit the computational cost of operating with belief functions in the Dempster-Shafer framework, the *plausibility transform*. Initially defined in terms of commonality values, the plausibility transform $\tilde{pl} : \mathscr{B} \to \mathscr{P}$, $b \mapsto \tilde{pl}[b]$ maps each belief function b onto the probability distribution $\tilde{pl}[b] = \tilde{pl}_b$ obtained by normalizing the plausibility values $pl_b(x)$[1] of the element of Θ:

$$\tilde{pl}_b(x) = \frac{pl_b(x)}{\sum_{y \in \Theta} pl_b(y)}. \tag{1}$$

We call the output (1) of the plausibility transform *relative plausibility of singletons* (r.pl.s.). Voorbraak proved that his (in our terminology) relative plausibility of singletons $\tilde{pl}_b$ is a perfect representative of b when combined with other probabilities $p \in \mathscr{P}$ through Dempster's rule $\oplus$: $\tilde{pl}_b \oplus p = b \oplus p \quad \forall p \in \mathscr{P}$.
Dually, a *relative belief transform* $\tilde{b} : \mathscr{B} \to \mathscr{P}$, $b \mapsto \tilde{b}[b]$ mapping each belief function to the corresponding *relative belief of singletons* (r.b.s.) $\tilde{b}[b] = \tilde{b}$ [6, 8, 12, 9]

$$\tilde{b}(x) = \frac{b(x)}{\sum_{y \in \Theta} b(y)} \tag{2}$$

can be defined. Unlike (1), however, (2) exists iff b assigns some mass to singleton focal sets: $\sum_{x \in \Theta} m_b(x) \neq 0$. The notion of relative belief transform (under the name of normalized belief of singletons) has first been proposed by Daniel [9]. Some preliminary analyses of the relative belief transform and its close relationship with the (relative) plausibility transform have been presented in [6, 8]. A detailed discussion of the geometrical properties of $\tilde{b}$ and $\tilde{pl}$ has been given in [7].

The epistemic family of probability transforms. Cobb and Shenoy [3] have argued in favor of the plausibility transform as a link between Shafer's theory of evidence (endowed with Dempster's rule) and Bayesian reasoning. They have proved

---

[1] With a harmless abuse of notation we denote the values of b.f.s and pl.f.s on a singleton x by $b(x), pl_b(x)$ rather than $b(\{x\}), pl_b(\{x\})$.

[2] that the plausibility transform commutes with Dempster's rule, and meets a number of additional properties which they claim "allow an integration of Bayesian and D-S reasoning that takes advantage of the efficiency in computation and decision-making provided by Bayesian calculus while retaining the flexibility in modeling evidence that underlies D-S reasoning":

$$b \oplus p = \tilde{pl}_b \oplus p \; \forall p; \quad \tilde{pl}_b[b_1 \oplus b_2] = \tilde{pl}_b[b_1] \oplus \tilde{pl}_b[b_2];$$
$$b \oplus b = b \vdash \tilde{pl}[b] \oplus \tilde{pl}[b] = \tilde{pl}[b].$$

On our side, we have proved [8] that a similar set of (dual) properties hold for the relative belief transform:

$$pl_b \oplus p = \tilde{b} \oplus p \; \forall p; \quad \tilde{b}[pl_{b_1} \oplus pl_{b_2}] = \tilde{b}[pl_{b_1}] \oplus \tilde{b}[pl_{b_2}];$$
$$pl_b \oplus pl_b = pl_b \vdash \tilde{b}[pl_b] \oplus \tilde{b}[pl_b] = \tilde{b}[pl_b],$$

where $pl_b \oplus$ denotes the extension of Dempster's rule to plausibility measures [8] (seen as pseudo belief functions, i.e., sum functions $pl_b(A) = \sum_{B \subseteq A} \mu_b(B)$ on 2^Θ whose Moebius transform $\mu_b(B)$ can be negative for some $B \subset \Theta$). This supports the existence of a family of probability transformations strongly linked to Shafer's interpretation of the theory of evidence via Dempster's rule, which includes relative belief and relative plausibility transforms, and which we call *epistemic* family, in opposition to the *affine* family of mappings which commute with affine combination [4] (a property that Smets calls "linearity" [15]).

Paper contribution and outline. The symmetry/duality between (relative) plausibility and belief is, unfortunately, broken, as the existence of the relative belief of singletons is subject to a strong condition. This stresses the issue of its applicability for, in practice, the situation in which the mass of all singletons is nil is common. However, in Section 2 we point out that relative belief is only a member of a class of *relative mass* transformations which generalize it, are computable even when relative belief is not, and can be interpreted as low-cost proxies for both plausibility and pignistic transforms (Section 3). We discuss their applicability as approximate transformations in two significant scenarios (Section 4).

2 Generalizing the relative belief transform

No matter its semantics and that of its sister plausibility transform, a serious issue with the relative belief of singletons is its applicability. In opposition to relative plausibility, $\tilde{b}$ does not exist for a large class of belief functions (those which assign no mass to singletons). Indeed, in many practical applications there is a bias towards some particular models which are the most exposed to the problem. For example, in "consonant" belief functions [14] at most one focal element is a singleton, therefore the vast majority of the useful information in the b.b.a. is contained in the non-singleton focal elements.

Relative belief is in fact only one element of an entire family of probability transformations. Indeed, $\tilde{b}$ can be thought of as the transform which, given a b.f. b:

1. retains the focal elements of size 1 only, yielding an unnormalized b.f.;
2. computes (indifferently) the latter's relative plausibility/pignistic transformation:

$$\tilde{b}(x) = \frac{\sum_{A\supseteq x,|A|=1} m_b(A)}{\sum_y \sum_{A\supseteq x,|A|=1} m_b(A)} = \frac{m_b(x)}{k_{m_b}} = \frac{\sum_{A\supseteq x,|A|=1} \frac{m_b(A)}{|A|}}{\sum_y \sum_{A\supseteq x,|A|=1} \frac{m_b(A)}{|A|}}.$$

Accordingly, a family of natural generalizations of the relative belief transform is obtained by, given an arbitrary b.f. b:

1. retaining the focal elements of size s only;
2. computing either the resulting relative plausibility ...
3. ... or the associated pignistic transformation.

Now, both alternatives 2) or 3) *yield the same probability distribution*. Indeed, the application of the relative plausibility transform yields: $p(x) = \frac{\sum_{A\supseteq\{x\}:|A|=s} m_b(A)}{\sum_{y\in\Theta}\sum_{A\supseteq\{y\}:|A|=s} m_b(A)} = \frac{\sum_{A\supseteq\{x\}:|A|=s} m_b(A)}{\sum_{A\subseteq\Theta:|A|=s} m_b(A)|A|} = \frac{\sum_{A\supseteq\{x\}:|A|=s} m_b(A)}{s\sum_{A\subseteq\Theta:|A|=s} m_b(A)}$, while applying the pignistic transform yields:

$$p(x) = \frac{\sum_{A\supseteq\{x\}:|A|=s} \frac{m_b(A)}{|A|}}{\sum_{y\in\Theta}\sum_{A\supseteq\{y\}:|A|=s} \frac{m_b(A)}{|A|}} = \frac{s\sum_{A\supseteq\{x\}:|A|=s} m_b(A)}{s\sum_{y\in\Theta}\sum_{A\supseteq\{y\}:|A|=s} m_b(A)}, \tag{3}$$

i.e., the same result. The following natural extension of the relative belief operator is therefore well defined.

Definition 1. Given any b.f. $b : 2^\Theta \to [0,1]$ with b.b.a. m_b, we call *relative mass transformation* of level s the transform $\tilde{M}_s[b]$ which maps b to the probability (3). We denote by $\tilde{m}_s$ the output of the relative mass transform of level s.

3 Approximation of pignistic and plausibility transform

It is easy too see that both relative plausibility of singletons and pignistic probability are *convex combinations of all the (n) relative mass probabilities* $\{\tilde{m}_s, s = 1,...,n\}$. Namely, let us we denote by:

$$k_{b,s} = \sum_{A\subseteq\Theta:|A|=s} m_b(A), \quad pl_b(x;s) = \sum_{A\supseteq\{x\}:|A|=s} m_b(A)$$

the total mass of focal elements of size s, and the contribution to the plausibility of x of size-s focal elements, respectively. Immediately: $\sum_y pl_b(y) = \sum_y \sum_{A\supseteq\{y\}} m_b(A) = \sum_{A\subseteq\Theta} m_b(A)|A| = \sum_{r=1}^n r(\sum_{A\subseteq\Theta,|A|=r} m_b(A)) = \sum_{r=1}^n r k_{b,r}$. Therefore, we obtain for the relative plausibility of singletons the following convex decomposition into relative mass probabilities $\tilde{m}_s$: $\tilde{pl}_b(x) =$

$$= \frac{pl_b(x)}{\sum_y pl_b(y)} = \frac{\sum_s pl_b(x;s)}{\sum_r rk_{b,r}} = \sum_s \frac{pl_b(x;s)}{\sum_r rk_{b,r}} = \sum_s \frac{pl_b(x;s)}{sk_{b,s}} \frac{sk_{b,s}}{\sum_r rk_{b,r}} = \sum_s \alpha_s \tilde{m}_s(x), \tag{4}$$

as $\tilde{m}_s(x) = \frac{pl_b(x;s)}{sk_{b,s}}$, with coefficients $\alpha_s = \frac{sk_{b,s}}{\sum_r rk_{b,r}} \propto sk_{b,s} = \sum_y pl_b(y;s)$ measuring for each level s the total plausibility contribution of the focal elements of size s.

In the case of the pignistic probability we get:

$$\begin{aligned} BetP[b](x) &= \sum_{A \supseteq \{x\}} \frac{m_b(A)}{|A|} = \sum_s \sum_{A \supseteq \{x\}, |A|=s} \frac{m_b(A)}{s} = \sum_s \frac{1}{s} \sum_{A \supseteq \{x\}, |A|=s} m_b(A) \\ &= \sum_s \frac{1}{s} pl_b(x;s) = \sum_s k_{b,s} \frac{pl_b(x;s)}{sk_{b,s}} = \sum_s k_{b,s} \tilde{m}_s(x), \end{aligned} \tag{5}$$

with coefficients $\beta_s = k_{b,s}$ measuring for each level s the mass contribution of the focal elements of size s.

Accordingly, the relative mass probabilities can be seen as basic components of both the pignistic and the plausibility transform, associated with the evidence carried by focal elements of a specific size.
As such transforms can be computed just by considering size-s focal elements, they can also be thought of as low-cost proxies for both relative plausibility and pignistic probability, since only the $\binom{n}{s}$ size-s focal elements (instead of the initial 2^n) have to be stored, while all the others can be dropped without further processing.

We can think of two natural criteria for such an approximation of $\tilde{pl}$, $BetP$ via the relative mass transforms.

- (C1) in the convex decompositions (4) and (5) associated with $\tilde{pl}$ and $BetP$, respectively, we retain the component s whose coefficient (α_s in the first case, β_s in the second) is the largest;
- (C2) we retain the component associated with the *minimal size* focal elements.

Clearly, the relative belief transformation coincides with the approximation produced by (C2) if $\sum_x m_b(x) \neq 0$. When the mass of singletons is nil, instead, the second criterion delivers a natural extension of the relative belief operator:

$$\tilde{b}^{ext}(x) \doteq \frac{\sum_{A \supseteq \{x\}:|A|=min} m_b(A)}{|A|_{min} \sum_{A \subseteq \Theta:|A|=min} m_b(A)}. \tag{6}$$

The two approximation criteria favor different aspects of the original belief function. (C1) focuses on the strength of the evidence carried by focal elements of equal size, by selecting those whose cardinality s is such that the total plausibility contribution of the focal elements of size s, $k_{b,s} = \sum_y pl_b(y;s)$, is the greatest. Note that, however, the optimal (C1) approximations of plausibility or pignistic transform are in principle quite distinct, as: $\hat{s}[\tilde{pl}] = \arg\max_s sk_{b,s}$, while $\hat{s}[BetP] = \arg\max_s k_{b,s}$. The best approximation for the pignistic probability will not necessarily be the best approximation of the relative plausibility of singletons. Criterion (C2) favors instead the *precision* of such pieces of evidence, measured by the size of the corresponding focal elements. Let us compare these two approaches in two simple scenarios.

4 Two scenarios

While C1 is (at least superficially) a sensible, rational principle (the selected proxy must be the greatest contributor to the actual classical probability transformation), C2 seems harder to justify. Why should one retain only the smallest focal elements, regardless their mass?
The attractive feature of the relative belief of singletons, among C2 approximations, is its simplicity: the original mass is directly re-distributed onto the singletons. What about the "extended" operator (6)?

4.1 Scenario 1

Consider a scenario in which we want to approximate the plausibility/pignistic transform of a b.f. $b : 2^\Theta \rightarrow [0,1]$, with b.b.a. $m_b(A) = m_b(B) = \varepsilon$, $|A| = |B| = 2$, and $m_b(\Theta) = 1 - 2\varepsilon \gg m_b(A)$ (Figure 1-left). Its relative plausibility of singletons is given by:

$$\begin{array}{ll}\tilde{pl}_b(x) \propto m_b(A) + m_b(\Theta), & \tilde{pl}_b(y) \propto m_b(A) + m_b(B) + m_b(\Theta), \\ \tilde{pl}_b(z) \propto m_b(B) + m_b(\Theta), & \tilde{pl}_b(w) \propto m_b(\Theta) \; \forall w \neq x, y, z.\end{array}$$

Its pignistic probability values are:

$$\begin{array}{ll}BetP(x) = \frac{m_b(A)}{2} + \frac{m_b(\Theta)}{n}, & BetP(y) = \frac{m_b(A)+m_b(B)}{2} + \frac{m_b(\Theta)}{n}, \\ BetP(z) = \frac{m_b(B)}{2} + \frac{m_b(\Theta)}{n}, & BetP(w) = \frac{m_b(\Theta)}{n} \forall w \neq x, y, z.\end{array}$$

Assuming $m_b(A) > m_b(B)$, both transformations have a profile as in Figure 1-right.

Fig. 1 Left: the original b.f. in the first scenario discussed in the text. Right: corresponding profile of both relative plausibility of singletons and pignistic probability.

Now, according to (C1), the best approximation (among all relative mass transormations) of both $\tilde{pl}_b$ and $BetP[b]$ is given by selecting the focal elements of size n, i.e., Θ, as the greatest contributor to both the convex sums (4) and (5).
However, it is easy to see that this yields as an approximation the average probability $\tilde{m}_1(w) = 1/n \; \forall w \in \Theta$, which carries no information at all. In particular, the fact that the available evidence supports to a limited extent the singletons x, y and z is completed discarded, and no decision is possible.
If, on the other hand, we operate according to the criterion (C2), we end up selecting the size-2 focal elements A and B. The resulting approximation is

$$\tilde{m}_2(x) \propto m_b(A), \; \tilde{m}_2(y) \propto m_b(A) + m_b(B), \; \tilde{m}_2(z) \propto m_b(B),$$

$\tilde{m}_2(w) = 0 \; \forall w \neq x, y, z$. This has the same profile as that of $\tilde{pl}_b$ or $BetP[b]$ (Figure 1-right): the decision made corresponds to that made based on $\tilde{pl}_b$ or $BetP[b]$.

In a decision-making sense, therefore, $\tilde{m}_2 = \tilde{b}^{ext}$ is the best approximation of both plausibility and pignistic transforms. We end up making the same decision, at a much lower (in general) computation cost.

4.2 Scenario 2

Consider however a second scenario, in which a b.f. has only two focal elements A and B, with $|A| > |B|$ and $m_b(A) \gg m_b(B)$ (Figure 2-left). Both relative plausibility

Fig. 2 Left: the b.f. of the second scenario. Right: corresponding profile of both relative plausibility of singletons and pignistic probability.

and pignistic probability have the following values:

$$\tilde{pl}_b(w) = BetP(w) \propto m_b(A) \; w \in A, \quad \tilde{pl}_b(w) = BetP(w) \propto m_b(B) \; w \in B,$$

and correspond to the profile of Figure 2-right.
In this second case, (C1) and (C2) generate the uniform probability on elements of A (as $m_b(A) \gg m_b(B)$) and the uniform probability on elements of B (as $|B| < |A|$), respectively. Therefore, it is (C1) that yields the best approximation of both plausibility and pignistic transforms in a decision-making perspective.

5 Conclusions

In this paper we tried and enrich our understanding of the family of epistemic transforms of belief functions. We showed that relative belief is only a member of a class of *relative mass* transformations which generalize it, are computable even when the mass of singletons is nil, and can be interpreted as low-cost proxies for both plausibility and pignistic transforms. We discussed their applicability as approximate transformations in two significant scenarios.

References

1. T. Burger, "Defining new approximations of belief functions by means of Dempster's combination," in *Proc. of BELIEF'10*.

2. B. Cobb and P. Shenoy, "A comparison of methods for transforming belief function models to probability models," in *Proc. of ECSQARU'03*, Aalborg, Denmark, pp. 255–266.
3. B. Cobb and P. Shenoy, "A comparison of Bayesian and belief function reasoning," *Information Systems Frontiers*, vol. 5, no. 4, pp. 345–358, 2003.
4. F. Cuzzolin, "Two new Bayesian approximations of belief functions based on convex geometry," *IEEE Transactions on Systems, Man, and Cybernetics - Part B*, vol. 37, no. 4, pp. 993–1008, 2007.
5. F. Cuzzolin, "A geometric approach to the theory of evidence," *IEEE Transactions on Systems, Man, and Cybernetics - Part C*, vol. 38, no. 4, pp. 522–534, 2008.
6. F. Cuzzolin, "Dual properties of the relative belief of singletons," in *Proc. of the Pacific Rim International Conference on AI*, 2008, pp. 78–90.
7. F. Cuzzolin, "Geometry of relative plausibility and relative belief of singletons" *Annals of Mathematics and Artificial Intelligence*, pp. 1–33, 2010.
8. F. Cuzzolin, "Semantics of the relative belief of singletons," in *Workshop on Uncertainty and Logic*, Kanazawa, Japan, 2008.
9. M. Daniel, "On transformations of belief functions to probabilities," *Int. J. of Intelligent Systems*, vol. 21, no. 3, pp. 261–282, 2006.
10. A. P. Dempster, "A generalization of Bayesian inference," *Journal of the Royal Statistical Society, Series B*, vol. 30, pp. 205–247, 1968.
11. J. Dezert and F. Smarandache, "A new probabilistic transformation of belief mass assignment," in *Proc. of the 11th International Conference of Information Fusion*, 2008, pp. 1–8.
12. R. Haenni, "Aggregating referee scores: an algebraic approach," in *COMSOC'08, 2nd International Workshop on Computational Social Choice*, 2008, pp. 277–288.
13. J. Lowrance, T. Garvey, and T. Strat, "A framework for evidential-reasoning systems," in *Proc. of the National Conference on Artificial Intelligence*, 1986, pp. 896–903.
14. G. Shafer, *A mathematical theory of evidence*. Princeton University Press, 1976.
15. P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66, no. 2, pp. 191–234, 1994.
16. P. Smets, "Decision making in the TBM: the necessity of the pignistic transformation," *IJAR*, vol. 38, no. 2, pp. 133–147, 2005.
17. J. Sudano, "Equivalence between belief theories and nave Bayesian fusion for systems with independent evidential data," in *Proc. of the 6th International Conference of Information Fusion*, 2003, vol. 2, pp. 1239–1243.
18. B. Tessem, "Approximations for efficient computation in the theory of evidence," *Artificial Intelligence*, vol. 61, no. 2, pp. 315–329, 1993.
19. F. Voorbraak, "A computationally efficient approximation of Dempster-Shafer theory," *International Journal on Man-Machine Studies*, vol. 30, pp. 525–536, 1989.