Game-theoretical semantics of epistemic probability transformations

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, while they are not consistent with the credal or probability-bound semantic of belief functions. We prove here, however, that these transforms can be given in this latter case an interesting rationale in terms of optimal strategies in a non-cooperative game.

1 Introduction

The theory of evidence (ToE) [21] 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*. Different operators have been proposed for the combination of two or more belief functions, starting from the orthogonal sum originally formulated by A. Dempster [15]. 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.

Belief functions possess a number of alternative semantics in terms of multi-valued mappings, random sets [19], inner measures [17], transferable beliefs [25] or hints [18]. In some of his papers [15], Dempster claimed that the mass $m_b(A)$

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

associated with a non-singleton event $A \subseteq \Theta$ could be understood as a "floating probability mass". This has originated a popular but controversial interpretation of belief functions b as convex sets $\mathscr{P}[b]$ of probabilities (often called *consistent* with b) determined by sets of lower and upper bounds on their probability values: $\mathscr{P}[b] \doteq \{p \in \mathscr{P} : b(A) \leq p(A) \leq pl_b(A)\ \forall A \subseteq \Theta\}$, where the plausibility function $pl_b : 2^\Theta \to [0,1]$, $pl_b(A) = 1 - b(A^c)$ carries the same evidence as b. In [22] Shafer disavowed any probability-bound interpretation, also criticized by Walley as incompatible with Dempster's rule of combination [31], a position later seconded by Dempster [14].

Probability transformation of belief functions. Nevertheless, the relation between belief and probability in the theory of evidence has been an important subject of study, and a number of papers have been published on the issue of probability transform [13]. A decision based approach to the problem is the foundation of Smets' "Transferable Belief Model" [25], in which belief functions are defined directly in terms of basis belief assignments ("credal" level), while decisions are made via the *pignistic probability* $BetP[b](x) = \sum_{A \supseteq \{x\}} \frac{m_b(A)}{|A|}$, generated by what he calls the *pignistic transform*: $BetP : \mathscr{B} \to \mathscr{P}$, $b \mapsto BetP[b]$. The pignistic probability is the result of a redistribution process in which the mass of each focal element A is re-assigned to all its elements $x \in A$ on an equal basis, and is perfectly compatible with the upper-lower probability semantics of b.f.s, as it is the center of mass of the polytope $\mathscr{P}[b]$ of consistent probabilities [4].

Other proposals have been recently brought forward by Dezert et al. [16], Burger [3], Sudano [27] and others, based on redistribution processes similar to that of the pignistic transform. In addition, two new Bayesian approximations of belief functions have been derived from purely geometric considerations [7] in the context of the geometric approach to the ToE [8], in which belief and probability measures are represented as points of a Cartesian space.

Relative plausibility and belief transforms. Originally developed by Voorbraak [29] as a probabilistic approximation intended to limit the computational cost of operating with belief functions in the Dempster-Shafer framework, the *plausibility transform* [5] has later been supported by Cobb and Shenoy in virtue of its commutativity properties with respect to Dempster's sum. 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 to the probability distribution $\tilde{pl}[b] = \tilde{pl}_b$ obtained by normalizing the plausibility values $pl_b(x)$[1] of the element of Θ: $\tilde{pl}_b(x) = \frac{pl_b(x)}{\sum_{y \in \Theta} pl_b(y)}$.

We call the output $\tilde{pl}_b$ of the plausibility transform *relative plausibility of singletons*. Voorbraak proved that his (in our terminology) relative plausibility of singletons is a perfect representative of b when combined with other probabilities $p \in \mathscr{P}$ through Dempster's rule $\oplus$: $\tilde{pl}_b \oplus p = b \oplus p$ for all $p \in \mathscr{P}$.

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* $\tilde{b}(x) = \frac{b(x)}{\sum_{y \in \Theta} b(y)}$ can be defined.

[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 $b(\{x\}), pl_b(\{x\})$.

The notion of relative belief transform (under the name of "normalized belief of singletons") has first been proposed by Daniel [13]. Some preliminary analyses of the relative belief transform and its close relationship with the (relative) plausibility transform have been presented in [9, 10]. A detailed discussion of the geometrical properties of $\tilde{b}$ and $\tilde{pl}$ has been given in [11]. In [10], in particular, the author has shown that plausibility and belief transforms both commute with Dempster's rule of combination, and meet a number of dual properties with respect to the orthogonal sum, therefore forming what we call the "epistemic" family of transforms. In opposition, an "affine" family can be defined which groups together those transforms which commute with affine combination, and fit in the probability-bound interpretation of belief functions.

Paper contribution. In this paper, instead, we point out that, even though they are not consistent with the credal set of probabilities dominated by the original belief function, plausibility and belief transforms can be provided in this interpretation with an interesting betting semantics within an adversarial game theory scenario [28]. In this scenario, inspired by Wald's minimax/maximin model [30], an opponent representing the uncertainty encoded by a b.f. is free to pick any probability function in the set determined by the latter: the decision maker's goal is to maximize their minimal expected reward (or minimize their maximal expected loss).

2 A game/utility theory interpretation

It can be proven that a probability distribution on Θ is consistent with a belief function b iff it is the result of a *redistribution process*, in which the mass of each focal element is shared between its elements in an arbitrary proportion [12]. However, neither the relative belief of singletons nor the relative plausibility of singletons (unlike Smets' pignistic function) are consistent in this sense: indeed, it is easy to prove that they are not the result of such a redistribution process [12]. Nevertheless, an interesting interpretation for them under the probability-bound semantic can be provided in a game/utility theory context [28, 26].

Strat's carnival wheel scenario. Consider the following scenario, inspired by Strat's expected utility approach to decision making with belief functions [26, 20]. In a country fair, by paying a fixed fee c, people get the chance to spin a carnival wheel divided into a number of sectors labeled, say, $\Theta = \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$. In return, they get an amount $r(x)$ which varies with the label $x \in \Theta$ of the sector that stops at the top, so that the gain or "utility" of each outcome for the player is $u(x) = r(x) - c$, while their loss is, dually, $l(x) = -u(x) = c - r(x)$.
The game amounts to a "lottery" (probability distribution), in which the probability of each outcome is proportional to the area covered on the wheel. People are asked to make a binary decision: to play/not to play. A rational behavior on the player's side consists on computing their expected utility $\sum_{x\in\Theta} u(x)p(x)$ and decide to play if the latter is positive: the decision, lacking any uncertainty, is trivial.

Cloaked carnival wheel. Strat therefore introduces a more challenging scenario, in which the fair's manager decides to make the game more interesting by covering part of the wheel. People are still asked whether they want to spin the wheel or not, knowing that the manager is allowed to rearrange the hidden sector of the wheel as they pleases (see Figure 1). Clearly, this new situation amounts to a set of possible

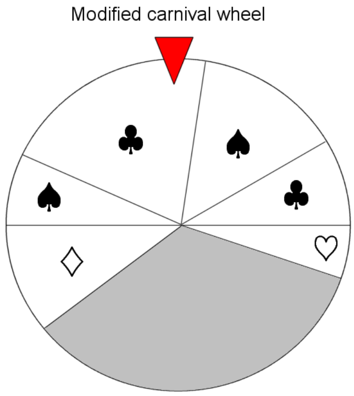

Fig. 1 The modified carnival wheel: part of the spinning wheel is cloaked.

lotteries which can be described as a belief function, in particular one in which the fraction of area associated with the hidden sector is assigned as mass to the whole decision space $\{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$. If additional (partial) information is provided, for instance that $\diamondsuit$ cannot appear in the hidden sector, different belief functions must be chosen instead. Regardless the particular belief function b (seen as a set of probabilities) at hand, the rule allowing the manager to pick an arbitrary distribution of outcomes in the hidden section mathematically translates into allowing them to choose *any* probability distribution $p \in \mathscr{P}[b]$ consistent with b in order to damage the player. Strat uses this situation as a way of introducing upper and lower bounds to the expected utility [28]

$$E(u) = \sum_{x \in \Theta} u(x) p(x),$$

of the player, induced by the upper and lower bounds to probabilities associated with the belief function describing the set of lotteries [26].

A modified carnival wheel scenario. Let us consider, instead, a modified scenario in which players are asked (after paying the usual fee c) *to bet on a single outcome* $x \in \Theta$. What is the expected utility of the player in this case? Clearly:

$$E(u) = \sum_{y \in \Theta} p(y) u'(y), \qquad u'(y) = \begin{cases} 0 & y \neq x \\ u(x) = r(x) - c & y = x \end{cases}$$

so that $E(u) = \sum_{y \in \Theta} p(y) u'(y) = p(x) u(x)$.
Suppose that the aim of the player is to play conservatively, and maximize their *worst case* expected utility $p(x)u(x)$, under the uncertainty given by the countermove by the fair's manager: which outcome (singleton) should they pick?

Wald's minimax model. This situation can be naturally described by *Wald's maximin model* [30], a non-probabilistic, robust decision making model in which the optimal decision is one whose worst outcome is at least as good as the worst outcome in any other case. Mathematically, it reads as follows:

$$f^* = \max_{a \in \mathscr{A}} \min_{s \in \mathscr{S}(a)} f(a,s) \tag{1}$$

where $\mathscr{A}$ denotes the set of alterative actions/decisions/strategies, $\mathscr{S}(a)$ denotes the set of states associated with action s, and $f(a,s)$ denotes the return of strategy a taking place in the state s. The model represents a 2-person game in which the max player plays first, making a move a: in response, the second (min) player selects the available state ($s \in \mathscr{S}(a)$) which minimizes the return for the first player.
Wald's model (1) represents a major simplification of the classic 2-person zero sum game [2], in which the two players decide without being aware of the other's choice, while in this case the players choose sequentially.

A minimax model of the carnival wheel, and relative beliefs. Clearly, our scenario can be described by a maximin model (1), in which: the set of possible actions corresponds to the set of outcomes of the lottery $\mathscr{A} = \Theta$; the set of possible states the second player can pick from does not depend on $a = x$, and is the set $\mathscr{S}(a) = \mathscr{S} = \mathscr{P}[b]$ of probability distributions consistent with b; and finally, the return is the player's expected utility $E(u) = p(x)u(x)$ under the constraint of having to pick a single outcome: $f(a,s) = f(x,p) = p(x)u(x)$ which is a function of the lottery outcome only. The problem may therefore be described as:

$$x_{maximin} = \arg\max_{x \in \Theta} \min_{p \in \mathscr{P}[b]} u(x)p(x). \tag{2}$$

Now, in the probability-bound interpretation of belief functions, the belief value of each singleton $x \in \Theta$ measures the minimal support x can receive from a distribution of the family associated with b: $b(x) = \min_{p \in \mathscr{P}[b]} p(x)$. Therefore

$$\begin{aligned} x_{maximin} &= \arg\max_{x \in \Theta} \min_{p \in \mathscr{P}[b]} u(x)p(x) = \arg\max_{x \in \Theta} \Big(u(x) \min_{p \in \mathscr{P}[b]} p(x) \Big) \\ &= \arg\max_{x \in \Theta} u(x)b(x) = \arg\max_{x \in \Theta} u(x)\tilde{b}(x) \end{aligned}$$

is the optimal decision for the player since, by normalizing $b(x)$ to obtain the relative belief of singletons, the maximal decision is obviously preserved.
If, in particular, the utility function is constant (i.e., no element of Θ can be preferred over the others), the best possible defensive strategy $x_{maximin}$ aimed at maximizing the minimal return of the possible outcomes is/are the peak(s) of the relative belief of singletons. In the example of Figure 1, as ♣ is the outcome which occupies the largest share of the visible part of the wheel, the safest bet (the one which guarantees the best expected return in the worst case) is indeed ♣.

Dual maximin model, and relative plausibilities. The dual *maximin* model describes the case in which the player moves first again, but this time to minimize the worst possible expected loss. In the modified carnival wheel scenario, once again, when people are asked to bet on a single outcome, their expected loss is $E(l) = l(x)p(x)$ so that:

$$x_{minimax} = \arg\min_{x\in\Theta} \max_{p\in\mathscr{P}[b]} l(x)p(x) = \arg\min_{x\in\Theta} l(x) \max_{p\in\mathscr{P}[b]} p(x) = \arg\min_{x\in\Theta} l(x)pl_b(x), \tag{3}$$

as $pl_b(x) = \max_{p\in\mathscr{P}[b]} p(x)$ measures the maximal possible support to x by a distribution consistent with b. Since $l(x) = c - r(x) = -u(x)$, and after noting that normalizing the plausibility of singletons does not alter the above optimization problem, the outcome/action which minimizes the maximal expected loss is:

$$x_{minimax} = \arg\min_{x\in\Theta} -u(x)\tilde{pl}_b(x) = \arg\max_{x\in\Theta} u(x)\tilde{pl}_b(x).$$

Once again, if in particular the loss (utility) function is constant, then the elements whose relative plausibility is maximal are the best possible defensive strategies aimed at minimizing the maximum possible loss.
In both the maximin and the minimax scenarios, relative belief and plausibility of singletons play a crucial role in determining the safest betting strategy in an adversarial game in which the decision maker has to minimize their maximal expected loss/maximize their minimal expected return under uncertainty representable as a belief function, interpreted as a set of lower/upper bounds to probability values.

The role of expected utility in pignistic transform. It can be useful to compare our scenario based on the maximin/minimax model with classical expected utility theory [28]. There, a decision maker can choose between a number of "lotteries" (probability distributions) $p_i(x)$, in order to maximize the expected return or utility $E(p_i) = \sum_x u(x)p_i(x)$ of the lottery. Here, the "lottery" is chosen by their opponent (given the available partial evidence), and the decision maker is left with betting on the safest strategy (element of Θ).

However, a look at how expected utilities are employed in the justification of Smets' pignistic transform provides a useful hint on a natural generalization of the proposed scenario. In [24], the author proves the necessity of the linearity axiom (and therefore of the pignistic transform) by maximizing the following expected utility (our notation), where $p = BetP$ is the pignistic function: $E[u] = \sum_{x\in\Theta} u(a,x)p(x)$. In this case, the set of possible actions (decision) $\mathscr{A}$ and the set Θ of possible outcomes of the problem are distinct, and the utility function is defined on $\mathscr{A} \times \Theta$.

A generalization of the proposed scenario. We can then wonder what happens if we generalize our scenario to the more general case in which the second player still impersonates the uncertainty on the lottery represented by a belief function, but the set of actions $\mathscr{A}$ is fully distinct from Θ, so that a utility function $u : \mathscr{A} \times \Theta \to \mathbb{R}^+$ can be defined. Let us focus on the *maximin* form, while forgetting the carnival wheel situation to move to a more abstract setting.
In this case, once again, the max player moves first and picks an action $\bar{a} \in \mathscr{A}$. This fixes a utility profile $u(\bar{a},x)$, $x \in \Theta$ for the elements of Θ: the first player now has a non-zero utility $u(\bar{a},x)$ for any possible outcome x of the problem, so that their expected utility is obviously given by $\sum_{x\in\Theta} u(\bar{a},x)p(x)$, which depends on the actual probability distribution describing the problem. The min player at this point selects

the admissible probability distribution $p \in \mathscr{P}[b]$ which minimizes the expected return of the max player. The overall model is in this more general case:

$$a_{maximin} = \arg\max_{a \in \mathscr{A}} \min_{p \in \mathscr{P}[b]} \Big(\sum_{x \in \Theta} u(a,x) p(x) \Big). \tag{4}$$

We can notice that, in this new situation, $\arg\max_{a \in \mathscr{A}} \min_{p \in \mathscr{P}[b]} \left(\sum_{x \in \Theta} u(a,x)p(x) \right) \neq \arg\max_{a \in \mathscr{A}} \left(\sum_{x \in \Theta} u(a,x) \min_{p \in \mathscr{P}[b]} p(x) \right) = \arg\max_{a \in \mathscr{A}} \left(\sum_{x \in \Theta} u(a,x) b(x) \right) = \arg\max_{a \in \mathscr{A}} \left(\sum_{x \in \Theta} u(a,x) \tilde{b}(x) \right)$ for the worst case probability distribution $p^*(x) = \arg\min_{p \in \mathscr{P}[b]} p(x)$ is, in general, different for each $x \in \Theta$, and we cannot simply swap the min and $\sum$ operators. As a consequence, the generalization of Wald's maximin/minimax model to the case in which the second player represents the uncertainty associated with a belief function (in the probability-bound interpretation), but actions/decisions are distinct from the outcomes of the problem is no more a function of belief and plausibility values on singletons, and cannot be solved by using only the knowledge encoded by relative plausibilities and beliefs of singletons. A deeper study of this and more general settings is in order.

3 Conclusions

Epistemic transforms commute with Dempster's rule but they are not consistent with the probability bound interpretation of belief functions. Nevertheless, in this paper we proposed an interesting, novel interpretation of relative belief and plausibility of singletons as tools to provide optimal conservative strategies in a maximin/minimax 2-person game scenario derived from Wald's model, in which a player has to optimize their minimal expected gain/maximal expected loss under epistemic uncertainty in the form of a belief function. The study of more general models will be the goal of further research in the near future.

References

1. P. Bogler, "Shafer-Dempster reasoning with applications to multisensor target identification systems," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 17, no. 6, pp. 968–977, 1987.
2. S. Bowles, *Microeconomics: Behavior, institutions, and evolution*, Princeton University Press, 2004.
3. T. Burger, "Defining new approximations of belief functions by means of Dempster's combination," in *Proc. of BELIEF*, Brest, France, 2010.
4. A. Chateauneuf and J. Jaffray, "Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion," *Mathematical Social Sciences*, vol. 17, pp. 263–283, 1989.
5. B. Cobb and P. Shenoy, "A comparison of Bayesian and belief function reasoning," *Information Systems Frontiers*, vol. 5, no. 4, pp. 345–358, 2003.
6. F. Cuzzolin, "Geometry of Dempster's rule of combination," *IEEE Trans. on Systems, Man and Cybernetics B*, vol. 34, no. 2, pp. 961–977, 2004.
7. F. Cuzzolin, "Two new Bayesian approximations of belief functions based on convex geometry," *IEEE Trans. on Systems, Man, and Cybernetics B*, vol. 37, no. 4, pp. 993–1008, 2007.

8. F. Cuzzolin, "A geometric approach to the theory of evidence," *IEEE Trans. on Systems, Man, and Cybernetics C*, vol. 38, no. 4, pp. 522–534, 2008.
9. F. Cuzzolin, "Dual properties of the relative belief of singletons," in *Proc. of PRICAI*, Hanoi, Vietnam, pp. 78–90, 2008.
10. F. Cuzzolin, "Semantics of the relative belief of singletons," in *Workshop on Uncertainty and Logic*, Kanazawa, Japan, 2008.
11. F. Cuzzolin, "The geometry of consonant belief functions: simplicial complexes of necessity measures," *Fuzzy Sets and Systems*, vol. 161, no. 10, pp. 1459–1479, 2010.
12. F. Cuzzolin, "On the relative belief transform," *International Journal of Approximate Reasoning*, 2012 (in press).
13. M. Daniel, "On transformations of belief functions to probabilities," *Int. J. of Intelligent Systems*, vol. 21, no. 3, pp. 261–282, 2006.
14. A. P. Dempster, "Lindley's paradox: Comment," *Journal of the American Statistical Association*, vol. 77, no. 378, pp. 339–341, 1982.
15. A. P. Dempster, "A generalization of Bayesian inference," in *Classic Works of the Dempster-Shafer Theory of Belief Functions*, 2008, pp. 73–104.
16. 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.
17. R. Fagin and J. Halpern, "Uncertainty, belief and probability," in *Proc. of IJCAI*, pp. 1161–1167, 1989.
18. J. Kohlas and P.-A. Monney, *A Mathematical Theory of Hints. An Approach to Dempster-Shafer Theory of Evidence*, Springer-Verlag, 1995.
19. H. Nguyen, "On random sets and belief functions," *Journal of Mathematical Analysis and Applications*, vol. 65, pp. 531–542, 1978.
20. J. Schubert, "On 'rho' in a decision-theoretic apparatus of Dempster-Shafer theory," *International Journal of Approximate Reasoning*, vol. 13, pp. 185–200, 1995.
21. G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.
22. G. Shafer, "Constructive probability," *Synthese*, vol. 48, pp. 309–370, 1981.
23. P. Shenoy, "No double counting semantics for conditional independence," Working Paper No. 307. School of Business, University of Kansas, 2005.
24. P. Smets, "Decision making in the TBM: the necessity of the pignistic transformation," *International Journal of Approximate Reasoning*, vol. 38, no. 2, pp. 133–147, 2005.
25. P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66, no. 2, pp. 191–234, 1994.
26. T. M. Strat, "Decision analysis using belief functions," *International Journal of Approximate Reasoning*, vol. 4, pp. 391–417, 1990.
27. J. Sudano, "Equivalence between belief theories and nave Bayesian fusion for systems with independent evidential data," in *Proc. of ICIF*, 2003, vol. 2, pp. 1239–1243.
28. J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 1944.
29. F. Voorbraak, "A computationally efficient approximation of Dempster-Shafer theory," *International Journal on Man-Machine Studies*, vol. 30, pp. 525–536, 1989.
30. A. Wald, *Statistical decision functions*, Wiley, New York, 1950.
31. P. Walley, "Belief function representations of statistical evidence," *The Annals of Statistics*, vol. 15, pp. 1439–1465, 1987.
32. H. Xu, Y.-T. Hsia, and P. Smets, "The transferable belief model for decision making in the valuation-based systems," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 26, pp. 698–707, 1996.
33. L. Zadeh, "A simple view of the Dempster-Shafer theory of evidence and its implications for the rule of combination," *AI Magazine*, vol. 7, no. 2, pp. 85–90, 1986.