Game-theoretical semantics of epistemic probability transformations

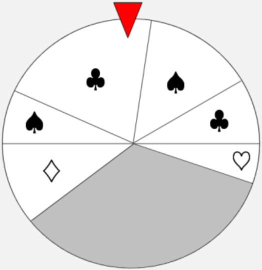
Semantics of epistemic transformations

Fabio Cuzzolin

Department of Computing and Communication Technologies
Oxford Brookes University

BELIEF 2012 - Compiegne, France
May 9-11th, 2012

Outline



Belief versus probability measures

- a **belief function** $b : 2^\Theta \rightarrow [0,1]$ is such that,

$$b(A) = \sum_{B \subseteq A} m(B)$$

where $m : 2^\Theta \rightarrow [0,1]$ is a basic belief (probability) assignment s.t.

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

- **Bayesian** belief functions: all focal elements are singletons
- are in 1-1 correspondence with (finite) probability measures

Mapping a belief function to a possibility

- the relation between belief and probability, in particular, has been an important subject of study in the theory of evidence
- **probability transform** of belief functions: an operator $pt : \mathcal{B} \rightarrow \mathcal{P}$, $b \mapsto pt[b]$ mapping belief measures onto probability distributions
- not necessarily “consistent” with original b.f., i.e., an element of the corresponding credal set
- they are **less computationally expensive** than b.f.s
- allow to **make decisions in the expected utility framework**
- a number of transforms proposed, either as efficient implementations of ToE or tools for decision making

Some literature on probability transforms

- very much not exhaustive!
- [*Lowrance* 86] "summarization" technique
- [*Smets* 88] pignistic transform
- [*Voorbraak* 89] plausibility transform
- [*Tessem* 93] m_{klx} approximation incorporating only the highest-valued focal elements
- [*Cobb&Shenoy* 03] plausibility transform and Dempster's rule
- [*Sudano* 03], [*Dezert* 07], [*Burger* 10] redistribution processes similar to that of the pignistic transform
- [*Daniel* 06] normalized belief of singletons
- [*Cuzzolin* 07] orthogonal projection and intersection probability
- [*Cuzzolin* 10] geometry of plausibility and belief transform

Plausibility transform

- other transforms have been proposed for efficient implementation of belief calculus
- plausibility transform [Voorbraak 89]
- maps a belief function to the **relative belief of singletons**:

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

- relative plausibility of singletons $\tilde{pl}_b$ is a perfect representative of b when combined with other probabilities $p \in \mathcal{P}$ by Dempster's rule $\oplus$
- meets a number of properties w.r.t. $\oplus$ [Coob&Shenoy 03]

Relative belief transform

- the *relative belief transform* $\tilde{b} : \mathcal{B} \rightarrow \mathcal{P}$, $b \mapsto \tilde{b}[b]$ ...
- ... maps each belief function to the corresponding *relative belief of singletons*:

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

- it exists iff b assigns some mass to singleton focal sets: $\sum_{x \in \Theta} m_b(x) \neq 0$
- first been proposed by Daniel in 2006
- its geometry and that of plausibility transform analyzed in [Cuzzolin 2010 AMAI]

Properties and duality

- a similar set of (dual) properties hold for the relative belief transform [Cuzzolin 12 IJAR]

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

- $pl_b \oplus$ denotes the extension of Dempster's rule to plausibility measures, seen as pseudo belief functions
- there exists a family of probability transformations strongly linked to Shafer's interpretation of the theory of evidence via Dempster's rule: the **epistemic family** of transformation
- in opposition to the **affine family** of mappings which commute with affine combination [Smets, Cuzzolin 07 SMCB]

Credal semantics and consistent transformations

- Belief functions possess a number of **alternative semantics**: multi-valued mappings [Dempster], random sets [Nguyen], inner measures [Fagin], transferable beliefs, hints [Kohlas]
- **credal** interpretation: belief functions b as convex sets $\mathcal{P}[b]$ of "consistent" probabilities

$$\mathcal{P}[b] \doteq \left\{ p \in \mathcal{P} : b(A) \leq p(A) \leq pl_b(A) \; \forall A \subseteq \Theta \right\},$$

- Shafer disavowed any probability-bound interpretation
- also criticized by Walley as incompatible with Dempster's rule of combination

Transformation in the TBM: the pignistic transform

- **pignistic probability**

$$BetP[b](x) = \sum_{A \supseteq \{x\}} \frac{m_b(A)}{|A|},$$

  the result of a *pignistic transform*
- its purpose is to allow decision making at the level of probabilities, typically in an expected utility framework
- 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
- it commutes with affine combination and is the center of mass of the credal set of consistent probabilities

Epistemic transformations are not consistent!

Theorem

The relative belief of singletons is not always consistent.

- counterexample: a b.f. on $\Theta = \{x_1, x_2, ..., x_n\}$ with b.b.a. $m_b(x_i) = k_{m_b}/n$ for all i, $m_b(\{x_1, x_2\}) = 1 - k_{m_b}$

Theorem

The relative plausibility of singletons is not always consistent.

- counterexample: a b.f. on $\Theta = \{x_1, x_2, x_3\}$ with b.b.a.

$$m_b(\{x_i\}^c) = \frac{k}{3} \; \forall i = 1, 2, 3, \quad m_b(\{x_1, x_2\}^c) = m_b(\{x_3\}) = 1 - k.$$

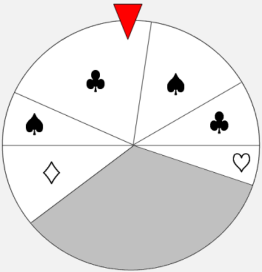
Consistency and redistribution process

- 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
- in the case of plausibility and belief transform no such redistribution exists
- still, an interpretation for them exists in the credal semantics
- a **utility interpretation in an adversarial game scenario**
- scenario, inspired by Strat's expected utility approach to decision making with belief functions [Strat'90]

Strat's carnival wheel scenario

- by paying a fixed fee c, people get the chance to spin a carnival wheel divided into a number of sectors labeled, say, $\Theta = \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$
- they get an amount $r(x)$ which varies with the label $x \in \Theta$ of the sector that stops at the top
- the gain or "utility" of each outcome for the player is $u(x) = r(x) - c$, their loss $l(x) = -u(x) = c - r(x)$
- this amounts to a "lottery" (distribution), in which the probability of each outcome is proportional to its area on the wheel
- a rational player's behavior: computing their expected utility $\sum_{x \in \Theta} u(x)p(x)$; play if the latter is positive
- lacking any uncertainty, decision is trivial

Cloaked carnival wheel



- the fair's manager **covers part of wheel**
- new situation: **set of possible lotteries, described as a belief function**
- area of hidden sector → mass of the whole decision space $\{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$
- the manager can pick *any* probability distribution $p \in \mathcal{P}[b]$ consistent with b to damage the player
- Strat introduces upper and lower bounds to the expected utility $E(u) = \sum_{x \in \Theta} u(x)p(x)$

A modified carnival wheel scenario

- modified scenario in which players are (after paying the usual fee c) **asked to bet on a single outcome** $x \in \Theta$
- expected utility of the player in this case: $E(u) = p(x)u(x)$
- 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 counter-move by the fair's manager
- which outcome (singleton) should they pick?

Wald's minimax model

- naturally described by **Wald's maximin model** [Wald'50]
- non-probabilistic, robust decision making model
- the optimal decision is one whose worst outcome is at least as good as the worst outcome in any other case:

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

- $\mathcal{A} \rightarrow$ set of alterative actions/decisions/strategies
  $\mathcal{S}(a) \rightarrow$ set of states associated with action s
  $f(a, s) \rightarrow$ return of strategy a taking place in the state s
- *2-person game in which the* max *player plays first, making a move a: in response, the second* (min) *player selects the available state* $(s \in \mathcal{S}(a))$ *which minimizes the return for the first player*

A minimax model of the wheel, and relative beliefs

- our scenario can be described by a maximin model (1) where
  - the set of possible actions is the set of lottery outcomes $\mathcal{A} = \Theta$
  - the set of possible states the second player can pick from is the set $\mathcal{S}(a) = \mathcal{S} = \mathcal{P}[b]$ of probability distributions consistent with b
  - the return is the player's expected utility $f(a, s) = f(x, p) = p(x)u(x)$, a function of the lottery outcome only
- problem is described as

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

- in the credal semantics of BFs $b(x) = \min_{p \in \mathcal{P}[b]} p(x)$ so that

$$x_{maximin} = \arg\max_{x \in \Theta} u(x)\tilde{b}(x)$$

Dual maximin model, and relative plausibilities

- dual *maximin* model: case in which the player moves first again, but this time to minimize the worst possible expected loss
- the model becomes:

$$x_{minimax} = \arg\min_{x \in \Theta} l(x) pl_b(x)$$

since $pl_b(x) = \max_{p \in \mathcal{P}[b]} p(x)$

- since $l(x) = c - r(x) = -u(x)$, and normalizing the plausibility of singletons does not alter the above optimization problem:

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

Conclusion!

- **in both the maximin and the minimax scenarios, relative belief and plausibility of singletons determine 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

- compare with classical expected utility theory [von Neumann'44]
- 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
- look at how expected utilities are employed in the justification of Smets' pignistic transform
- Smets proves the necessity of the pignistic transform by maximizing $E[u] = \sum_{x \in \Theta} u(a, x)p(x)$
- the set of possible actions $\mathcal{A}$ and the set Θ of possible outcomes are distinct, and the utility function is defined on $\mathcal{A} \times \Theta$

A generalization of the proposed scenario

- what happens if we generalize our scenario to the more general case in which **the set of actions $\mathcal{A}$ is fully distinct from Θ**?
- focus on the *maximin* form, move to a more abstract setting
- the max player moves first and picks an action $\bar{a} \in \mathcal{A}$: their expected utility is $\sum_{x \in \Theta} u(\bar{a}, x) p(x)$
- the min player at this point selects the $p \in \mathcal{P}[b]$ which minimizes the expected return of the max player
- overall model is in this **more general case**:

$$a_{maximin} = \arg\max_{a \in \mathcal{A}} \min_{p \in \mathcal{P}[b]} \left(\sum_{x \in \Theta} u(a,x) p(x) \right) \neq \arg\max_{a \in \mathcal{A}} \left(\sum_{x \in \Theta} u(a,x) \tilde{b}(x) \right)$$

- we cannot simply swap the min and $\sum$ operators
- no more a function of beliefs and plausibilities of singletons

Conclusions

- epistemic transforms commute with Dempster's rule but they are **not consistent** with the probability bound interpretation of belief functions
- we proposed a novel interpretation of relative belief and plausibility of singletons in the credal interpretation
- within a game theoretical framework
- as tools to provide **optimal conservative strategies in a maximin/minimax 2-person game scenario** derived from Wald's model
- the generalization to distinct action sets will likely require to move away from them