Semantics of the relative belief of singletons

Fabio Cuzzolin

INRIA Rhône-Alpes
Institut National de Recherche en Informatique et en Automatique

International Workshop on Interval/Probabilistic Uncertainty and Non-Classical Logics (UncLog'08)
Japan Advanced Institute of Science and Technology (JAIST)
Ishikawa, Japan, March $24^{th} - 28^{th}$ 2008

Outline

Belief and probability measures

- basic belief assignment and **belief functions**
  - basic belief assignment $m : 2^\Theta \to [0,1]$ such that

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

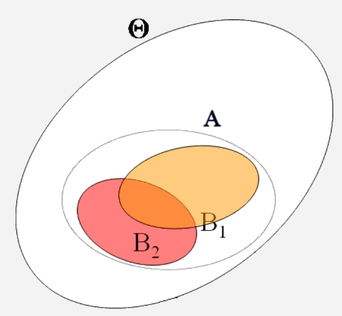
  - *belief function* $b : 2^\Theta \to [0,1]$,

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

- probability functions or Bayesian b.f.s: $m_b(A) = 0, |A| > 1$
- plausibility function: $pl_b(A) = 1 - b(A^c) = \sum_{B \cap A} m_b(B)$

Belief and probability measures
Example

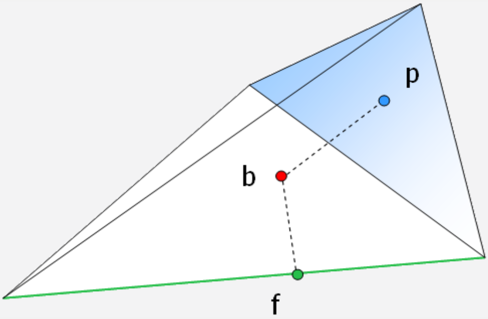
- **focal elements**, sets with non-zero mass (B_1, B_2);



- belief of A: $b(A) = m(B_1) + m(B_2)$;

Probabilistic approximation

- problem: **finding the probability measure** p **which is the closest** to a given belief function b



- several different approaches

Existing approaches

- **pignistic function** [Smets]

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

  is the center of mass of all probabilities dominating b: $p(A) \geq b(A) \; \forall A$
- **orthogonal projection** of b onto the probability simplex $\mathcal{P}$ [Cuzzolin]
- **intersection probability** [Cuzzolin]: assigns to each element the same fraction of the uncertainty given by a probability interval

$$p[b](x) = m_b(x) + \beta[b](pl_b(x) - m_b(x))$$

Relative plausibility of singletons

- **relative plausibility** of singletons

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

- assigns to each element x its normalized plausibility
- proven to be **equivalent** to the original b.f. b when combined with a probability [Voorbraak]

$$b \oplus p = pl_b \oplus b \quad \forall p \in \mathcal{P}$$

- **commutes** with respect to Dempster's rule [Cobb and Shenoy]

Relative belief of singletons

- similar expression in which pl_b is replaced by b

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

- assigns to each element its normalized belief value
- exists if and only if some mass is assigned to singletons

$$\sum_{x\in\Theta} m_b(x) \neq 0.$$

Semantics of relative plausibility

- focal elements can be interpreted as **constraints** to the mass distribution of the (unknown) true probability
- in this sense relative plausibility has the following interpretation:
    - for each singleton $x \in \Theta$ the most *optimistic* hypothesis in which the mass of all $A \supseteq \{x\}$ focuses on x is considered, yielding $\{pl_b(x), x \in \Theta\}$;
    - this assumption, however, is contradictory as it is supposed to hold for all singletons (many of which belong to the same higher-size events);
    - nevertheless, the obtained values are normalized to yield a Bayesian belief function.
- **optimistic** estimate of support to x

Dual semantics of relative belief

- **conservative** estimate of support to x:
- for each singleton $x \in \Theta$ the most *pessimistic* hypothesis in which only the mass of $\{x\}$ itself actually focuses on x is considered, yielding $\{b(x) = m_b(x), x \in \Theta\}$;
- this assumption is also *contradictory*, as the mass of all higher-size events is not assigned to any singletons;
- the obtained values are again *normalized* to produce a Bayesian belief function
- duality also holds in a different sense

Pseudo belief functions

- a belief function is a function on 2^Θ whose Moebius inverse m_b (b.p.a.) meets the positivity axiom
- all functions $\varsigma : 2^\Theta \to \mathbb{R}$ admit Moebius inverse m_ς s.t.

$$\varsigma(A) = \sum_{B \subseteq A} m_\varsigma(B)$$

  where $m_\varsigma(B) \not\geq 0 \; \forall B \subseteq \Theta$ [Aigner]
- functions ς which meet the normalization constraint

$$\sum_{\emptyset \subsetneq A \subseteq \Theta} m_\varsigma(A) = 1$$

  are then natural extensions of belief functions
- we call them **pseudo belief functions**

Plausibilities are also pseudo belief functions

- plausibility functions are themselves pseudo belief functions
- their Moebius inverse

$$\mu_b(A) \doteq \sum_{B \subseteq A} (-1)^{|A \setminus B|} pl_b(B) = (-1)^{|A|+1} \sum_{B \supseteq A} m_b(B), \qquad A \neq \emptyset$$

is called **basic plausibility assignment**, $\mu_b(\emptyset) = 0$

Theorem

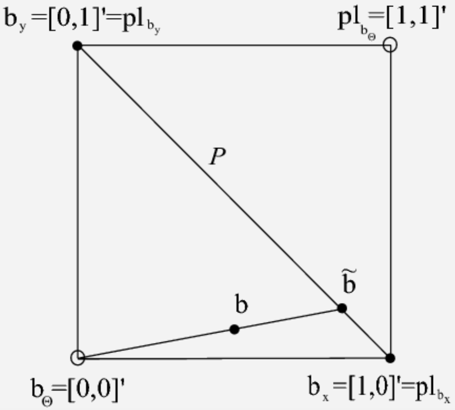
$m_b(x) = \sum_{A \supseteq \{x\}} \mu_b(A)$.

Duality

- the belief of singletons $b(x)$ is nothing but **the plausibility of singletons of pl_b interpreted as a pseudo belief function**:
$$b(x) = pl_{pl_b}(x)$$
  - each p.b.f. admits a (pseudo) plausibility function:
$$pl_\varsigma(A) = \sum_{B \cap A \neq \emptyset} m_\varsigma(B);$$
  - but for the above class of p.b.f. $\varsigma = pl_b$, so that
$$pl_{pl_b}(A) = \sum_{B \cap A \neq \emptyset} \mu_b(B)$$
  - when applied to singletons this yields
$$pl_{pl_b}(x) = \sum_{B \ni x} \mu_b(B) = m_b(x)$$

Geometry in the binary case

- case of belief functions on $\Theta = \{x, y\}$: $b = [b(x), b(y)]'$

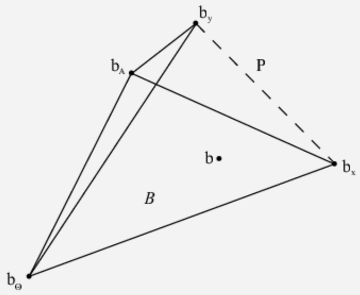


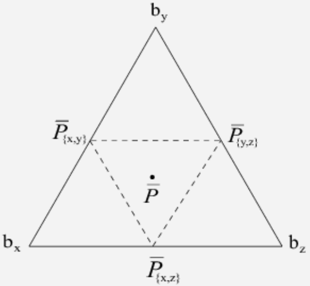
The existence condition

- unlike other Bayesian approximations, $\tilde{b}$ does not always exists
- only when the original b.f. b does assign some mass to elements of Θ
- however, in the binary case the only b.f. which does not admit rel.bel. is the vacuous one b_Θ: $m_{b_\Theta}(\Theta) = 1$
- $m_{b_\Theta}(x) = m_{b_\Theta}(y) = 0$ so that $\sum_x m_{b_\Theta}(x) = 0$ and $\tilde{b}_\Theta$ does not exist
- this is true in general, i.e. **the set of belief measures for which $\tilde{b}$ does not is lower-dimensional**

Singularity of the zero-mass condition

- the space of all belief functions is a *simplex*
- the probability simplex $\mathcal{P}$ is a face of $\mathcal{B}$





- in general terms *the whole* probability simplex $\mathcal{P}$ can host such approximations
- when $\{b : \sum_x m_b(x) = 0\}$ Bayesian approximations span only a proper subset of the probability simplex

Relative belief as approximation to other approaches

- relative belief is clearly more suitable for certain classes of belief functions
- one my guess this class corresponds to belief functions which are ”almost” probabilities

Theorem

For quasi-Bayesian b.f.s all Bayesian approximations converge:

$$\lim_{k_{m_b}\to 1} BetP[b] = \lim_{k_{m_b}\to 1} \tilde{pl}_b = \lim_{k_{m_b}\to 1} \tilde{b}.$$

- result states that relative belief can be used to approximate other Bayesian transformations when not too far from $\mathcal{P}$
- but computational cost of $\tilde{b}$ is much lower!

Ternary example

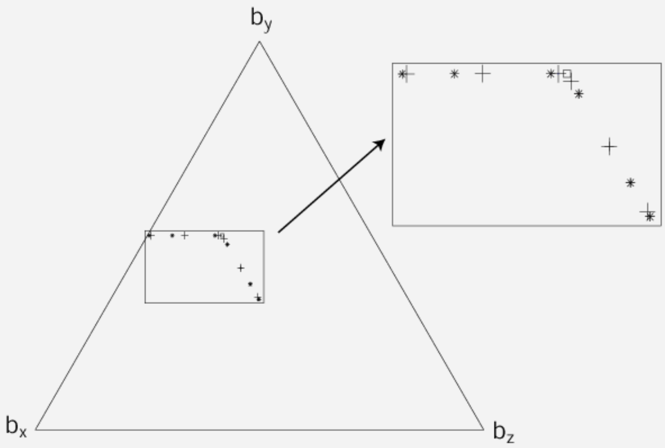


Figure: Convergence of pignistic function and relative plausibility to the relative belief in the ternary frame $\Theta = \{x, y, z\}$. Sample locations in the probability simplex of $\tilde{b}$ (square), BetP[b] (stars) and $\tilde{pl}_b$ (crosses) for $k_{m_b} = 0.95$, $k_{m_b} = 0.5$, $k_{m_b} = 0.05$

Ternary example

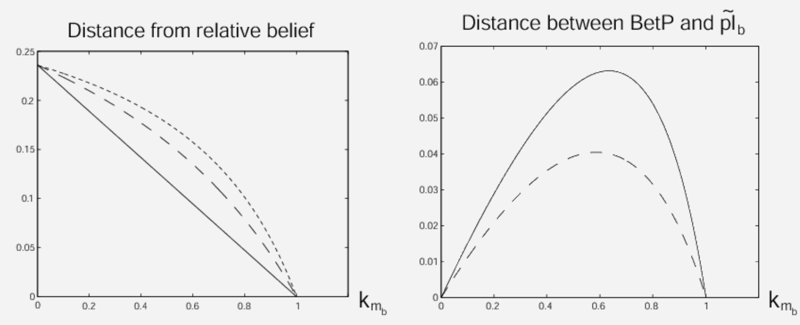


Figure: Left: distance from $\tilde{b}$ of $BetP[b]$ (solid line) and $\tilde{pl}_b$ (dotted line: case a; dashed line: case b) as a function of k_{m_b}. Right: Corresponding distance between $BetP$ and $\tilde{pl}_b$ (solid line: case a; dashed line: case b).

Dual properties

... of relative belief and plausibility

Proposition

The relative belief operator ***commutes*** *with respect to Dempster's combination of plausibility functions, namely*

$$\tilde{b}[pl_1 \oplus pl_2] = \tilde{b}[pl_1] \oplus \tilde{b}[pl_2].$$

The relative belief of singletons $\tilde{b}$ ***represents perfectly*** *the corresponding plausibility function* pl_b *when combined with any probability through (extended) Dempster's rule:*

$$\tilde{b} \oplus p = pl_b \oplus p$$

for each Bayesian belief function $p \in \mathcal{P}$.

Dual properties

... of relative belief and plausibility

Proposition

*If pl_b is **idempotent** with respect to Dempster's rule, i.e. $pl_b \oplus pl_b = pl_b$, then $\tilde{b}[pl_b]$ is itself **idempotent**:*

$$\tilde{b}[pl_b] \oplus \tilde{b}[pl_b] = \tilde{b}[pl_b].$$

If $\exists x \in \Theta$ such that $b(x) > b(y)\ \forall y \neq x, y \in \Theta$, then

$$\tilde{b}[pl_b^\infty](x) = 1, \quad \tilde{b}[pl_b^\infty](y) = 0\ \forall y \neq x$$

where pl_b^∞ denotes the infinite limit of the combination of pl_b with itself.

Conclusions
... and further thoughts

- relative belief of singletons as **dual** to relative plausibility
- **conservative estimate** of the evidence supporting elements of the frame
- dual interpretation as relative **plausibility of a plausibility** function
- **low-cost** approximation for other Bayesian transformations when the b.f. is quasi-Bayesian
- **dual properties** with respect to the rule of combination