

Advice-Giving and Scoring-Rule-Based Arguments for Probabilism

Lara Buchak & Branden Fitelson

Department of Philosophy
University of California–Berkeley

buchak@berkeley.edu
branden@fitelson.org

- Arguments for probabilism can be read as delivering some *non-specific* “advice” to incoherent agents: ***Be Coherent!***
- Standard arguments for probabilism are all of the form:
 - An agent S has a non-probabilistic degree of belief function b iff (\Leftrightarrow) S has some “bad” property B (presumably, *in virtue of* the fact that their b has a “bad” formal property F).
- These *arguments* rest on *Theorems* (\Rightarrow) and *Converse Theorems* (\Leftarrow): b is non-Pr $\Leftrightarrow b$ has formal property F .
 - **Dutch Book Arguments.** B is *susceptibility to sure monetary loss* (in a certain betting set-up), and F is the formal role played by non-Pr b 's in the DBT and the Converse DBT.
 - **Representation Theorem Arguments.** B is *having preferences that violate some of Savage's axioms* (and/or *being unrepresentable as an expected utility maximizer*), and F is the formal role played by non-Pr b 's in the RT.
- To the extent that we have reasons to avoid these B 's, these arguments provide reasons (not) to have a(n) (in)coherent b .
- Scoring Rule arguments *seem* to yield *more specific* “advice”.

- By “Scoring Rule Arguments” for probabilism, we mean:
 - B is *being in an “accuracy-dominated” epistemic state* (the precise meaning of B is what we’re trying to get clear on).
 - F is *inadmissibility*: \exists a set of statements P and a d.o.b.f. b^* , which (a) is coherent on P , and (b) *s-dominates* b on P , where s is some “good” scoring rule that is adopted by S (our main goal here is to get clearer on the implications of F).
- 👉 Since SRTs deliver *specific families* of dominating coherent d.o.b. functions b^* on P , one may be tempted to read them as yielding (epistemic) reasons to *adopt some such b^* on P* .
- That’s a *more specific* sort of reason than a mere reason to adopt *some coherent b' on P* , as in traditional arguments.
 - Analogy (to which we’ll return): having reason to adopt *some β'* that is logically consistent on P vs having reason to adopt *some β^** in a *specific family* of β ’s that are consistent on P .
- The rest of this talk is a *cautionary tale* about putting scoring rule arguments to this more specific sort of use.
- Our tale involves 3 theorems, an example, and an analogy.

- We’ll consider *logically omniscient* agents S , with languages \mathcal{L} & *total* d.o.b. functions b that are (i) $\mathcal{L} \mapsto [0, 1]$, (ii) assign the same value to \mathcal{L} -equivalent p ’s, (iii) $b(\top) = 1$, $b(\perp) = 0$.
- We’ll use “ b is incoherent on a set P (of statements in \mathcal{L})” to mean that the values b assigns to the elements of P are not compatible with any probability function over \mathcal{L} .
- **Theorem 1.** The following two properties of a set P of propositions (of \mathcal{L}) are *logically equivalent*:
 1. Some d.o.b. functions b (as above) are incoherent on P .
 2. Some (dogmatic) *full belief* functions $\beta \in \{0, 1\}$ are logically inconsistent on P (i.e., some $\beta(P)$ ’s are not compatible with any *truth-value assignment* over the entire language \mathcal{L}).

[[We won’t actually *use* this theorem — but, later, we will discuss an analogy between d.o.b. functions b that are incoherent on P and (dogmatic) full belief functions β that are inconsistent on P .]]
- **Theorem 2.** b is non-probabilistic on \mathcal{L} iff b is incoherent on (i.e., does not sum to 1 on) some *partition* of \mathcal{L} .
- We *will* make use of Theorem 2 — actually, we’ll *misuse* it...

- One might use a scoring rule theorem (*via*, say, Theorem 2) to try to generate “specific advice” for an incoherent S .
 - Suppose S (who has a language \mathcal{L} , and who adopts some “good” scoring rule s) has a non-probabilistic b .
 - Then, by Theorem 2, there will be a partition P of \mathcal{L} on which b is incoherent (*i.e.*, on which b does not sum to one).
 - So, by your favorite partition-based SRT, there will exist some b^* , which (a) is coherent on P , and (b) s -dominates b on P . Conversely, no such b^* will be dominated in this way.
 - At this point, you might be tempted to conclude that S has reason to adopt *some* such b^* on *partition* P .
 - The following example suggests that this may be hasty.
- Consider an agent S with a 2-atomic-sentence (X, Y) \mathcal{L} , and a d.o.b. function b on \mathcal{L} , which satisfies these six constraints:

$b(X \& Y) = \frac{1}{10}$	$b(X \& \sim Y) = \frac{2}{5}$	$b(\sim X \& Y) = \frac{1}{5}$
$b(\sim X \& \sim Y) = \frac{3}{10}$	$b(X) = \frac{1}{2}$	$b(\sim X) = \frac{2}{5}$

- Note that b is coherent on the *partition of state descriptions* of \mathcal{L} , but b is incoherent on *two other partitions* of \mathcal{L} .

- b is incoherent on (*exactly*—see next slide) these 2 P 's of \mathcal{L} :
 - (P_1) $\{X, \sim X\}$
 - (P_2) $\{X \& Y, X \& \sim Y, \sim X\}$
- So, scoring rule theorems will entail *both* of the following:
 - (1) There exists a b_1^* which (a) is coherent on P_1 , and (b) s -dominates b on P_1 (wrt S 's “good” scoring rule s). And, (conversely) no such b_1^* will be dominated in this way.
 - (2) There exists a b_2^* which (a) is coherent on P_2 , and (b) s -dominates b on P_2 (wrt S 's “good” scoring rule s). And, (conversely) no such b_2^* will be dominated in this way.
- Thus, if we applied our specific-advice-generating argument (above) to both (1) and (2), then we would conclude *both*:
 - S has reason to adopt some b_1^* on P_1 .
 - S has reason to adopt some b_2^* on P_2 .
- 👉 But: **Theorem 3.** Assuming that S adopts the Brier score (we will assume this from now on), it is *impossible* for S to *both* adopt some b_1^* on P_1 *and* adopt some b_2^* on P_2 .
- So, our naïve use of scoring rule arguments has led to the generation of *confusing* “specific advice” for *this* agent S .

- $\mathbb{A} \stackrel{\text{def}}{=} \text{the set of all (16) propositions of } \mathcal{L}. \text{ Finest-grained look:}$

p	$b_{\mathbb{A}}(p)$	$b'_{\mathbb{A}}(p)$	$b_2^*(p)$	$b_1^*(p)$	$b_{\mathbb{A}}^\dagger(p)$
$\sim X \& \sim Y$	3/10	3/10			23/80
$X \& \sim Y$	2/5	2/5	13/30		33/80
$X \& Y$	1/10	1/10	2/15		9/80
$\sim X \& Y$	1/5	1/5			3/16
$\sim Y$	7/10	7/10			7/10
$(\sim X \& \sim Y) \vee (X \& Y)$	2/5	2/5			2/5
$\sim X$	2/5	1/2	13/30	9/20	19/40
X	1/2	1/2		11/20	21/40
$(X \& \sim Y) \vee (\sim X \& Y)$	3/5	3/5			3/5
Y	3/10	3/10			3/10
$X \vee \sim Y$	4/5	4/5			13/16
$\sim X \vee \sim Y$	9/10	9/10			71/80
$\sim X \vee Y$	3/5	3/5			47/80
$X \vee Y$	7/10	7/10			57/80
$X \vee \sim X$	1	1	1	1	1
$X \& \sim X$	0	0	0	0	0

- $b_{\mathbb{A}}$ is the completion of b to \mathbb{A} that is incoherent on exactly P_1, P_2 .
- $b'_{\mathbb{A}}$ (a Pr on \mathbb{A}) seems “close” to $b_{\mathbb{A}}$, but does **not** Brier-dominate $b_{\mathbb{A}}$ (on \mathbb{A}).
- b_i^* are Euclidean-closest Pr’s to b (on P_i) that Brier-dominate b (on P_i).
- $b_{\mathbb{A}}^\dagger$ is the Euclidean-closest Pr to $b_{\mathbb{A}}$ (on \mathbb{A}) that Brier-dominates $b_{\mathbb{A}}$ (on \mathbb{A}).

- Consider an agent S (with $\beta_{\mathbb{A}}$) in a (global) preface case.
 - Does S have reason to adopt *some* $\beta'_{\mathbb{A}}$ that’s consistent on \mathbb{A} ?
 - Does S have reason to adopt *some* $\beta_{\mathbb{A}}^*$ from a *specific family* of β ’s that are logically consistent on \mathbb{A} ?
- Perhaps S has some reason to adopt *some* consistent $\beta'_{\mathbb{A}}$, since that’s the only way for S to avoid being such that *she knows a priori that some of her beliefs are false* (“bad” B).
- But, it *doesn’t* seem that S need have any reason to adopt a $\beta_{\mathbb{A}}^*$ from any *specific family* of consistent β ’s.
- Now, return to an agent S with a $b_{\mathbb{A}}$ that is incoherent on \mathbb{A} .
 - Does S have reason to adopt *some* $b'_{\mathbb{A}}$ that’s coherent on \mathbb{A} ?
 - Does S have reason to adopt a $b_{\mathbb{A}}^*$ from a *specific family* of b ’s that are coherent on \mathbb{A} (*viz.*, those which *s-dominate* $b_{\mathbb{A}}$ on \mathbb{A})?
- SRAs *seem* to justify affirmative answers to *both* questions.

👉 Proposal: (a) speak *only* of “incoherence **on** \mathbb{A} ” (which is *logically equivalent* to *inadmissibility*), and (b) use SRAs — **on** \mathbb{A} — to provide “specific advice”. [[Show visualization.]]