Belief and Credence
(The view from naïve epistemic utility theory)

Branden Fitelson

Philosophy @ Rutgers
&
MCMP @ LMU
branden@fitelson.org

- We assume that our agent has a credence function $b(\cdot)$, which is *probabilistic*. Probabilism for $b(\cdot)$ can *itself* be motivated *via* EUT [25]. But, this is *common ground* here.
- We assume that our agent takes exactly one of three qualitative attitudes (B, D, S) toward each member of a finite agenda $\mathcal{A}$ of (classical, possible worlds) propositions.
- We do *not* assume that these qualitative judgments can be *reduced* to $b(\cdot)$. But, we will use $b(\cdot)$ to derive a *rational coherence constraint* for qualitative judgment sets $\mathbf{B}$ (on $\mathcal{A}$).
- This derivation requires both the agent's credence function $b(\cdot)$ and their *epistemic utility function* [11, 18, 22] $u(\cdot)$. Following Easwaran [3, 5], we assume our agent cares *only* about whether their qualitative judgments are *accurate*.
- Specifically, our agent attaches some *positive* utility ($\mathfrak{r}$) with making an *accurate* judgment, and some *negative* utility ($-\mathfrak{w}$) with making an *inaccurate* judgment (where $\mathfrak{w} \geq \mathfrak{r} > 0$).

- Because suspensions are neither accurate nor inaccurate, our agent will attach *zero* epistemic utility to suspensions $S(p)$, independently of the truth-value of p.
- Thus, we have the following piecewise definition of $u(\cdot, w)$.

$$u(B(p), w) \stackrel{\text{def}}{=} \begin{cases} -\mathfrak{w} & \text{if } p \text{ is false at } w \\ \mathfrak{r} & \text{if } p \text{ is true at } w \end{cases}$$

$$u(D(p), w) \stackrel{\text{def}}{=} \begin{cases} \mathfrak{r} & \text{if } p \text{ is false at } w \\ -\mathfrak{w} & \text{if } p \text{ is true at } w \end{cases}$$

$$u(S(p), w) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } p \text{ is false at } w \\ 0 & \text{if } p \text{ is true at } w \end{cases}$$

- With this *accuracy-centered* epistemic utility function in hand, we can derive a naïve EUT coherence requirement.

- To do so, we'll also need a *decision-theoretic principle*.
- Applications of EUT to grounding probabilism as a (synchronic) requirement for $b(\cdot)$ typically appeal to a *non-dominance* (in epistemic utility) principle [14, 26, 25].
- But, some authors apply an *expected epistemic utility maximization* (or *expected inaccuracy minimization*) principle to derive rational requirements [17, 10, 4, 24].

> **Coherence**. An agent's belief set $\mathbf{B}$ over an agenda $\mathcal{A}$ should, from the point of view of their own credence function $b(\cdot)$, *maximize expected epistemic utility* (or *minimize expected inaccuracy*). That is, $\mathbf{B}$ should maximize
>
> $$EEU(\mathbf{B}, b) \stackrel{\text{def}}{=} \sum_{p \in \mathcal{A}} \sum_{w \in W} b(w) \cdot u(\mathbf{B}(p), w)$$
>
> where $\mathbf{B}(p)$ is the agent's attitude toward p, and $W \stackrel{\text{def}}{=} \bigcup \mathcal{A}$.

- For now, we assume "*act-state independence*": $\mathbf{B}(p)$ and p are *b-independent* [9, 2, 1, 15]. We'll return to this issue.

- The consequences of **Coherence** are rather simple and intuitive. It is straightforward to prove the following result.

Theorem ([3]). An agent with credence function $b(\cdot)$ and qualitative judgment set $\mathbf{B}$ over agenda $\mathcal{A}$ satisfies **Coherence** *if and only if* for all $p \in \mathcal{A}$

$$B(p) \in \mathbf{B} \text{ iff } b(p) > \frac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}},$$

$$D(p) \in \mathbf{B} \text{ iff } b(p) < 1 - \frac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}},$$

$$S(p) \in \mathbf{B} \text{ iff } b(p) \in \left[1 - \frac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}}, \frac{\mathfrak{w}}{\mathfrak{r}+\mathfrak{w}}\right].$$

☞ In other words, **Coherence** *entails Lockean representability*, where the Lockean thresholds are determined by the way the agent (relatively) values accuracy *vs.* inaccuracy.

- This provides an elegant, EUT-based explanation of why Lockean representability is a rational requirement for agents with *both* credences *and* qualitative attitudes.
- Next, I will explain when **Coherence** entails *consistency*.

- Suppose our (naïve) agent has a belief set $\mathbf{B}_n$ on a *minimal inconsistent* agenda of size n (*e.g.*, $(n-1)$–ticket lottery).

Theorem ([5]). For all $n \geq 2$ and any probability function $\Pr(\cdot)$, the $\Pr(\cdot)$-Lockean-representability of $\mathbf{B}_n$ (with threshold t) *entails* deductive consistency of $\mathbf{B}_n$ *iff* $t \geq \frac{n-1}{n}$.

- If we combine this with Easwaran's **Coherence** theorem, we get the following result, regarding the conditions under which the **Coherence** of $\mathbf{B}_n$ *entails* the *consistency* of $\mathbf{B}_n$.

Theorem. For all $n \geq 2$, an agent with an accuracy-centered utility function u, a credence function $b(\cdot)$, and a belief set $\mathbf{B}_n$, the **Coherence** of $\mathbf{B}_n$ entails the consistency of $\mathbf{B}_n$ *iff*

$$(\dagger) \qquad \mathfrak{w} \geq (n-1) \cdot \mathfrak{r}.$$

☞ Insisting that **Coherence** *implies consistency* (wrt $\mathbf{B}_n$) *requires (naïve) agents to disvalue inaccuracy at least* $(n-1)$ *times as much as they value accuracy.*

- Of course, there will be *some* agents with epistemic utility functions u, which *do* satisfy (†). But, it is very odd (from a traditional Bayesian perspective) to *mandate* that such an agent's epistemic utility function *must* satisfy (†).
- Assuming MEEU is *sufficient* for epistemic rationality, this is precisely what we would be doing to such agents, if we were to impose deductive consistency as a rational requirement. Clearly, this would be an unacceptable consequence.
- For example, in Lottery Paradox cases, we can make n as large as we like. And, the larger we make n, the stronger (and more implausible) the constraint (†) becomes.
- This is not to say that there won't be *some* MEEU-agents for whom consistency *is* a rational requirement, for *some* $\mathbf{B}_n$'s. But, $\mathbf{B}_n$-consistency won't be a *universal* MEEU-requirement.
- In other words, consistency *outstrips* the MEEU-theory of epistemic rationality. Leitgeb [16] defends an alternative.

- According to Hannes's Stability Theory [16], a rational agent with credence function b (over a set of possible worlds W) believes a proposition p, *viz.*, $B(p)$, iff $b(p \mid y) > t$, for all $y \in \mathcal{Y}$, where $\mathcal{Y} = \{y \mid b(y) > 0 \text{ and } \neg B(\neg p)\}$.
- As Hannes explains, his theory will require that "Stable" rational agents have *consistent* (and *closed*) belief sets (*e.g.*, let $\mathbf{B}_n$ be a belief set over a minimal inconsistent set of $n > 3$ propositions in an $(n-1)$-ticket Lottery Paradox).
- So, by our argument above, Stability Theory must *outstrip* MEEU-theory, which does *not* require consistency of $\mathbf{B}_n$ (at least, this is not required for *every* MEEU-rational agent).
- Next, I'll discuss some features of the Stability Theory (ST), with an eye toward (a) bringing out some of its distinctive properties, and (b) bridging the gap between MEEU and ST.
- I'll use a simple guiding example to illustrate just how differently MEEU and ST can behave (even in simple cases).

- My guiding example will involve a set $W = \{w_1, w_2, w_3, w_4\}$ containing four possible worlds. We can think of the example as involving a language with two atomic sentences $\{X, Y\}$, so that the worlds correspond to *state descriptions*.
- The relevant underlying Boolean algebra will contain 16 propositions. This allows us to visualize the example using (stochastic) truth-tables representing the entire algebra.
- The example involves two (rational) agents: S_1 is an MEEU-agent and S_2 is an ST-agent. S_1's belief state $\mathbf{B}_1$ is determined by her credence function b_1 *and* her u. S_2's belief state $\mathbf{B}_2$ is determined by her credence function b_2.
- In order to ensure a fair comparison, we will suppose that both agents have a $1/2$-threshold for rational belief.
- For S_1, this means her u is such that $\mathfrak{r} = \mathfrak{w}$. For S_2, this means $t = 1/2$ in her criterion for stable belief (*i.e.*, S_2 believes q just in case q is p*-stable*, relative to b_2).

- S_1 and S_2 *share the same credence function* $b_1 = b_2 = b$. But, they have *very different belief states* $\mathbf{B}_1$ and $\mathbf{B}_2$ [23]. The following table depicts b, $\mathbf{B}_1$ and $\mathbf{B}_2$ (on the *contingent* p's).

w's	p	b	$\mathbf{B}_1$	$\mathbf{B}_2$
$\{w_1\}$	$\neg X \wedge \neg Y$	0.5	S	S
$\{w_2\}$	$X \wedge \neg Y$	0.25	D	S
$\{w_3\}$	$X \wedge Y$	0.125	D	S
$\{w_4\}$	$\neg X \wedge Y$	0.125	D	S
$\{w_1, w_2\}$	$\neg Y$	0.75	B	S
$\{w_1, w_3\}$	$X \equiv Y$	0.625	B	S
$\{w_1, w_4\}$	$\neg X$	0.625	B	S
$\{w_2, w_3\}$	X	0.375	D	S
$\{w_2, w_4\}$	$X \not\equiv Y$	0.375	D	S
$\{w_1, w_4\}$	Y	0.25	D	S
$\{w_1, w_2, w_3\}$	$X \vee \neg Y$	0.875	B	S
$\{w_1, w_2, w_4\}$	$\neg X \vee \neg Y$	0.875	B	S
$\{w_1, w_3, w_4\}$	$\neg X \vee Y$	0.75	B	S
$\{w_2, w_3, w_4\}$	$X \vee Y$	0.5	S	S

- There are (arbitrarily) small perturbations b' of b, which (a) do not alter the $1/2$-credence p's, (b) *lower the credence of* $\neg X \vee \neg Y$, but (c) *make it rational for* S_2 *to believe* $\neg X \vee \neg Y$.

w's	p	b	b'	$\mathbf{B}_1 = \mathbf{B}_1'$	$\mathbf{B}_2$	$\mathbf{B}_2'$
$\{w_1\}$	$\neg X \wedge \neg Y$	0.5	0.5	S	S	S
$\{w_2\}$	$X \wedge \neg Y$	0.25	0.2366	D	S	S
$\{w_3\}$	$X \wedge Y$	0.125	0.1295	D	S	D
$\{w_4\}$	$\neg X \wedge Y$	0.125	0.1339	D	S	S
$\{w_1, w_2\}$	$\neg Y$	0.75	0.7366	B	S	S
$\{w_1, w_3\}$	$X \equiv Y$	0.625	0.6295	B	S	S
$\{w_1, w_4\}$	$\neg X$	0.625	0.6339	B	S	S
$\{w_2, w_3\}$	X	0.375	0.3660	D	S	S
$\{w_2, w_4\}$	$X \not\equiv Y$	0.375	0.3705	D	S	S
$\{w_1, w_4\}$	Y	0.25	0.2634	D	S	S
$\{w_1, w_2, w_3\}$	$X \vee \neg Y$	0.875	0.8661	B	S	S
$\{w_1, w_2, w_4\}$	$\neg X \vee \neg Y$	0.875	0.8705	B	S	B
$\{w_1, w_3, w_4\}$	$\neg X \vee Y$	0.75	0.7634	B	S	S
$\{w_2, w_3, w_4\}$	$X \vee Y$	0.5	0.5	S	S	S

- This example brings out just how different MEEU-theory and Stability Theory are. Note, also, that the MEEU belief set ($\mathbf{B}_1$) is *consistent* (although, it is *not closed*, since $\neg B(\neg X \,\&\, \neg Y)$).
- Is there a way to bridge this gap between MEEU and ST? *I.e.*, is there some way to understand what ST requires (over-and-above MEEU), from an EUT perspective?
- Here's a conjecture regarding one possible way of getting to something the resembles ST, using the machinery of EUT.

  Conjecture. Let $\mathcal{Y}$ be any set of W-propositions (with nonzero b-credence). If a belief set $\mathbf{B}$ (on $\mathcal{A}$) maximizes

  $$EEU_{\mathcal{Y}}(\mathbf{B}, b) \overset{\text{def}}{=} \sum_{p \in \mathcal{A}} \sum_{w \in W} b(w \mid y) \cdot u(\mathbf{B}(p), w)$$

  for all $y \in \mathcal{Y}$, then $\mathbf{B}$ is *resiliently* Lockean representable by $b(\cdot \mid y)$, *for each* $y \in \mathcal{Y}$, with threshold $t = \frac{\mathfrak{w}}{\mathfrak{r} + \mathfrak{w}}$.

- If this conjecture is true, then "Stability Theory" emerges from "*resilient* expected epistemic utility maximization."

- Thus, from a naïve EUT-perspective, *either* ST imposes strong constraints on the u-functions of rational agents, *or* ST requires a much stronger, "resilient" kind of MEEU.
- Is "resilient" MEEU a plausible rational requirement? In the practical case, it seems clear that such a requirement would be too demanding (in general). Many actions we take to be rational would be rendered irrational by such an account.
- Why think the epistemic case is any different? Simply insisting that deductive cogency is a requirement of epistemic rationality is not a very illuminating answer here (especially in light of the results and examples above).
- Is there an independent (epistemic-value-theoretic) argument that *more than* MEEU is required for epistemic rationality (and, specifically, that "stability" is required)?
- Is there some *alternative* unified theory of (both practical and epistemic) rationality that undergirds "stability"?

- A requirement on rational belief (or rational action) is *partition-invariant* (PI) iff its prescriptions do not depend on how the underlying space of possibilities is partitioned.
- In the case of practical rationality (*viz.*, rational action), many philosophers endorse (PI) as a *desideratum* [12, 6, 7, 19, 13].
- Savage's theory [27] and standard causal decision theories [8, 28, 20, 29] are partition-*dependent*. This has led various authors [12, 6, 13] to endorse evidential decision theories.
- We defined **Coherence** "Savage–style," and we assumed *act-state independence* (ASI) to ensure (PI). For our present examples (*e.g.*, Lotteries) this is OK. *But*, see [9, 2, 15].[1]

☞ Lin & Kelly [21] show: *any* non-trivial, Lockean coherence constraint that entails *deductively cogency must be* partition *dependent — even in Lottery cases* (*i.e.*, *even if* ASI obtains).

[1] More generally, **Coherence** will satisfy (PI) if u satisfies following, for all partitions $\{X_i\}$ of W: $(\forall X_i)\,[u(\mathbf{B}(p), X_i) = \sum_{w \in W} b(w \mid X_i) \cdot u(\mathbf{B}(p), w)]$.

[1] M. Caie, *Rational Probabilistic Incoherence*, *Philosophical Review*, 2013.

[2] J. Carr, *Epistemic Utility Theory and the Aim of Belief*, 2014.

[3] K. Easwaran, *Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probability*, manuscript, September 2014.

[4] ______, *Expected Accuracy Supports Conditionalization—and Conglomerability and Reflection*, *Philosophy of Science*, 2013.

[5] K. Easwaran and B. Fitelson, *Accuracy, Coherence, and Evidence*, to appear in *Oxford Studies in Epistemology V*, T. Szabo-Gendler and J. Hawthorne (*eds.*).

[6] E. Eells, *Rational Decision and Causality*, CUP, 1982.

[7] ______, *Levi's "The Wrong Box"*, *Journal of Philosophy*, 1985.

[8] A. Gibbard & W. Harper. *Counterfactuals and two kinds of expected utility*, 1981.

[9] H. Greaves, *Epistemic Decision Theory*, *Mind*, 2013.

[10] H. Greaves and D. Wallace, *Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility*, *Mind*, 2006.

[11] C. Hempel, *Deductive-Nomological vs. Statistical Explanation*, 1962.

[12] R. Jeffrey, *The Logic of Decision*, Chicago, 1990.

[13] J. Joyce, *The Foundations of Causal Decision Theory*, CUP, 1999.

[14] ______, *Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief*, in F. Huber & C. Schmidt-Petri (*eds.*), *Degrees of Belief*, 2009.

[15] J. Konek and B. Levinstein, *The Foundations of Epistemic Decision Theory*, 2014.

[16] H. Leitgeb, *The Stability Theory of Belief*, *Philosophical Review*, 2014.

[17] H. Leitgeb and R. Pettigrew, *An Objective Justification of Bayesianism I & II*, *Philosophy of Science*, 2010.

[18] I. Levi, *Gambling with Truth*, 1967.

[19] ______, *The Wrong Box*, *Journal of Philosophy*, 1985.

[20] D. Lewis, *Causal decision theory*, *Australasian Journal of Philosophy*, 1981.

[21] H. Lin and K. Kelly, *A Geo-logical Solution to the Lottery Paradox, with Applications to Non-monotonic Logic*, 2012.

[22] P. Maher, *Betting on Theories*, 1993.

[23] D. Makinson, *Remarks on the Stability Theory of Belief*, 2014.

[24] G. Oddie, *Conditionalization, cogency, and cognitive value*, *British Journal for the Philosophy of Science*, 1997.

[25] R. Pettigrew, *Epistemic Utility Arguments for Probabilism*, *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), URL = <http://plato.stanford.edu/archives/win2011/entries/epistemic-utility/>.

[26] J. Predd, R. Seringer, E. Loeb, D. Osherson, H.V. Poor and S., Kulkarni, *Probabilistic coherence and proper scoring rules*, *IEEE Transactions*, 2009.

[27] L. Savage, *The Foundations of Statistics*, Dover, 1972.

[28] B. Skyrms, *Causal decision theory*, *Journal of Philosophy*, 1982.

[29] J.H. Sobel, *Taking Chances: Essays on Rational Choice*, 1994.