We assume our agent has a credence function $b(\cdot)$, which is probabilistic. Probabilism for $b(\cdot)$ can itself be motivated via EUT [26]. Here, I will take this to be common ground.

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 $B$ (on $\mathcal{A}$).

This derivation requires both the agent’s credence function $b(\cdot)$ and their epistemic utility function $[12, 19, 23] u(\cdot)$.

Following Easwaran [4, 6], we assume our agent cares only about whether their qualitative judgments are accurate.

Specifically, our agent attaches some positive utility ($r$) with making an accurate judgment, and some negative utility ($-w$) with making an inaccurate judgment (where $w \geq r > 0$).

Because suspensions are neither accurate nor inaccurate (per se), 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 $u(\cdot, w)$.

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

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

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

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

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

**Theorem** ([4]). An agent with credence function $b(\cdot)$ and qualitative judgment set $B$ over agenda $\mathcal{A}$ satisfies Coherence if and only if for all $p \in \mathcal{A}$

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

$$D(p) \in B \iff b(p) < \frac{r}{r + w},$$

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

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 \( B_n \) on a minimal inconsistent agenda of size \( n \) (e.g., \( (n - 1) \)-ticket lottery).

**Theorem.** For all \( n \geq 2 \) and any probability function \( \Pr(\cdot) \), the \( \Pr(\cdot) \)-Lockean-representability of \( B_n \) (with threshold \( t \)) entails deductive consistency of \( B_n \) 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 \( B_n \) entails the consistency of \( B_n \).

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

\[
\forall w \geq (n - 1) \cdot r.
\]

Insisting that Coherence implies consistency (wrt \( B_n \)) requires (naïve) agents to disvalue inaccuracy at least \( (n - 1) \) times as much as they value accuracy.


As Leitgeb explains, his theory will require that rational agents have consistent (and closed \( \vdash \) cogent) belief sets (even belief sets \( B_n \) in \( (n - 1) \)-ticket Lottery Paradoxes).

So, by our argument above, Stability Theory must outstrip MEEU-theory, which does not require consistency of \( B_n \) (at least, this is not required for every MEEU-rational agent).

Leitgeb’s theory has several (prima facie) odd consequences. I will focus on one problematic feature: the violation of partition-invariance. As we’ll see, Leitgeb’s theory is partition-sensitive in a particularly troubling way.

- See Extras for some other (prima facie) odd consequences.

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.

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 require that such an agent’s epistemic utility function must satisfy (†).

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.

\( \therefore \) \( B_n \)-consistency won’t be a universal MEEU-requirement.

In other words, consistency outstrips the MEEU-theory of epistemic rationality. Leitgeb [17] defends an alternative.

According to Leitgeb’s Stability Theory [17], a rational agent with credence function \( b \) (over a set of possible worlds \( W \)) believes a proposition \( p \), viz., \( B(p) \) iff \( p \) is entailed by some proposition \( B_w \) that is \( p \)-stable, where this is defined as:

\( p \)-stability. Given a probability model \( (W, b(\cdot)) \), a proposition \( x \in W \) is \( p \)-stable iff \( b(x | y) \geq 1/2, \) for all \( y \in W \) such that \( x \& y \neq \bot \) and \( b(y) > 0 \).

In the case of practical rationality (viz., rational action), many philosophers endorse (PI) as a desideratum [13, 7, 8, 20, 14].

Savage’s theory [28] and standard causal decision theories [9, 30, 21, 31] are partition-dependent. This has led various authors [13, 7, 14] 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 [10, 2, 16].\(^1\)

Lin & Kelly [22] show: adding cogency to any non-trivial probabilistic acceptance rule for belief will entail partition sensitivity — even in Lottery cases (i.e., even if ASI obtains).

And, Schurz [29] shows that all cogent Lockean theories (i.e., all Stability theories) must violate an even more plausible invariance constraint, which he calls Independence.

\(^1\)More generally, Coherence will satisfy (PI) if \( u \) satisfies following, for all partitions \( \{X_i\} \) of \( W \):
\[
\forall x_i \in X_i \left[ u(B(p), x_i) = \sum_{w \in W} b(w | X_i) \cdot u(B(p), w) \right].
\]
Stability Theory vs. MEEU

As Dorling explains, Bayesian confirmation theory (BCT) implies the following two verdicts regarding this example (pace what Quine & Duhem might have said about the case).

1. $E$ weakly disconfirms $T$
   $$b(T \mid E) = 0.897308 \lessapprox 0.9 = b(T)$$

2. $E$ strongly disconfirms $H$
   $$b(H \mid E) = 0.00299103 \lessapprox 0.59806 = b(H)$$

Leitgeb offers a qualitative analysis of Dorling’s example (in ST), in which the agent starts off (prior to learning $E$) believing $H$, $T$, and $\neg E$. Then, after learning $E$, the agent comes to disbelieve $H$, while believing both $E$ and $T$.

Crucial to Leitgeb’s analysis is his choice of $B_W = \{w_1\}$, which is the strongest $p$-stable set, relative to $\langle W, b(\cdot) \rangle$.

While Leitgeb’s qualitative analysis of the Dorling example is interesting, it is threatened by the non-Independence of Stability Theory. Allow me to explain, via a similar example.

The two BCT verdicts [(1), (2)] above remain exactly the same in this new model $\langle W’, b’(\cdot) \rangle$, since $C$ is $b’$-independent.

But, the strongest $p$-stable proposition relative to this new model $\langle W’, b’(\cdot) \rangle$ is now $B_W’ = \{w’_1, w’_2, w’_9, w’_{10}\}$, which does not entail $H$. So, Leitgeb’s ST-analysis is undermined.

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 $B$ (on $\mathcal{A}$) maximizes

$$EEU_\mathcal{Y}(B, b) \equiv \sum_{p \in \mathcal{A}} \sum_{w \in W} b(w \mid y) \cdot u(B(p), w)$$

for all $y \in \mathcal{Y}$, then $B$ is resiliently Lockean representable by $b’(\cdot \mid y)$, for each $y \in \mathcal{Y}$, with threshold $t = \frac{w}{1-w}$.

If this conjecture is true, then “Stability Theory” emerges from “resilient” expected epistemic utility maximization.”
• \( S_1 \) and \( S_2 \) share the same credence function \( b_1 = b_2 = b \).

But, they have very different belief states \( B_1 \) and \( B_2 \) [24]. The following table depicts \( b, B_1 \) and \( B_2 \) on the contingent \( p \)’s.

<table>
<thead>
<tr>
<th>( w )'s</th>
<th>( p )</th>
<th>( b )</th>
<th>( B_1 ) (MEEU(_{1/2} ))</th>
<th>( B_2 ) (ST(_{1/2} ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>{( w_1 )}</td>
<td>( \neg X \land \neg Y )</td>
<td>0.5</td>
<td>( S )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_2 )}</td>
<td>( X \land \neg Y )</td>
<td>0.25</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_3 )}</td>
<td>( \neg X \land Y )</td>
<td>0.125</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_4 )}</td>
<td>( \neg Y )</td>
<td>0.75</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_2 )}</td>
<td>( X \equiv Y )</td>
<td>0.625</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_3 )}</td>
<td>( \neg X )</td>
<td>0.625</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_2, w_3 )}</td>
<td>( X )</td>
<td>0.375</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_2, w_4 )}</td>
<td>( X \equiv Y )</td>
<td>0.375</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_4 )}</td>
<td>( Y )</td>
<td>0.25</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_2, w_3 )}</td>
<td>( X \lor \neg Y )</td>
<td>0.875</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_2, w_4 )}</td>
<td>( \neg X \lor \neg Y )</td>
<td>0.875</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_3, w_4 )}</td>
<td>( \neg X \lor Y )</td>
<td>0.75</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_2, w_3, w_4 )}</td>
<td>( X \lor Y )</td>
<td>0.5</td>
<td>( S )</td>
<td>( S )</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>( w )'s</th>
<th>( p )</th>
<th>( b )</th>
<th>( b' )</th>
<th>( B_1 = B_1' )</th>
<th>( B_2 = B_2' )</th>
</tr>
</thead>
<tbody>
<tr>
<td>{( w_1 )}</td>
<td>( \neg X \land \neg Y )</td>
<td>0.5</td>
<td>0.5</td>
<td>( S )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_2 )}</td>
<td>( X \land \neg Y )</td>
<td>0.25</td>
<td>0.2366</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_3 )}</td>
<td>( X \lor Y )</td>
<td>0.125</td>
<td>0.1295</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_4 )}</td>
<td>( \neg X \lor \neg Y )</td>
<td>0.125</td>
<td>0.1339</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_2 )}</td>
<td>( \neg Y )</td>
<td>0.75</td>
<td>0.7236</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_3 )}</td>
<td>( X \lor \neg Y )</td>
<td>0.625</td>
<td>0.6295</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_4 )}</td>
<td>( \neg X \lor \neg Y )</td>
<td>0.625</td>
<td>0.6339</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_2, w_3 )}</td>
<td>( X \lor \neg Y )</td>
<td>0.375</td>
<td>0.3660</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_2, w_4 )}</td>
<td>( Y )</td>
<td>0.25</td>
<td>0.3705</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_4 )}</td>
<td>( X \lor \neg Y )</td>
<td>0.25</td>
<td>0.2634</td>
<td>( D )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_2, w_3 )}</td>
<td>( \neg X \lor \neg Y )</td>
<td>0.875</td>
<td>0.8661</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_2, w_4 )}</td>
<td>( \neg X \lor \neg Y )</td>
<td>0.875</td>
<td>0.8705</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_1, w_3, w_4 )}</td>
<td>( \neg X \lor Y )</td>
<td>0.75</td>
<td>0.7634</td>
<td>( B )</td>
<td>( S )</td>
</tr>
<tr>
<td>{( w_2, w_3, w_4 )}</td>
<td>( X \lor \neg Y )</td>
<td>0.5</td>
<td>0.5</td>
<td>( S )</td>
<td>( S )</td>
</tr>
</tbody>
</table>