Imprecise decision theories for Lockeans

Hana Kalpak · Stockholm University · hana.kalpak@philosophy.su.se

Stockholm University

TWO CONCEPTIONS OF BELIEF

Categorical belief models treat doxastic attitudes as all-or-nothing. Having considered whether p, you will do one of the following:

believe p	take p to be the case	$p \in B$
disbelieve p	take p not to be the case	$\neg p \in B$
suspend judgment on p	neither believe nor disbelieve p	$p, \neg p \notin B$

Areas: Traditional epistemology, epistemic logic, AGM belief revision...

Graded belief models treat doxastic attitudes as coming in degrees. Having considered whether p, you assign it a **credence** $c(p) \in [0,1]$, representing your degree of confidence that p is the case.

Areas: Subjective probability theory, decision theory, all things "Bayesian"...

THE LOCKEAN THESIS

If both attitude types exist, we would expect rationality to constrain how they combine. But how?

The Lockean thesis (LT). There is some threshold value $r : 0 < r < 1$, such that for any proposition p, rationality permits you to

- **believe** p iff $c(p) \geq r$,
- **disbelieve** p iff $c(p) \leq (1-r)$, and
- **suspend judgment** on p iff $(1-r) \leq c(p) \leq r$.

LT does not make sense if credences are calculated from credal sets. Let $r = .5$, $c(p) = [.4, .6]$, and LT forbids each categorical attitude towards p.

INTRODUCING IMPRECISION

A nice result from epistemic decision theory (e.g. Easwaran, 2016; Dorst 2019): LT follows if rational agents **maximize the expected utility** (accuracy) of their categorical beliefs.

Q: Can we generalize the Lockean thesis to imprecise credences through (MEU-compatible) rules for imprecise decision making?

A: Not through the rules I consider here (E-admissibility, Γ-maximin, Maximality). They yield agenda-dependent norms for epistemic rationality, while (canonical) LT is agenda-independent.

REPRESENTING BELIEFS

Define an **agenda** $\mathcal{A}$ as a finite set of propositions, closed at least under $\neg$. Intuitively, this is the set of propositions of which an agent, at some time-point, is aware.

Your **categorical beliefs** are represented by your **belief set** $B \subset \mathcal{A}$, interpreted as indicated above.

Your **graded beliefs** are represented by your **credal set** C: a closed, convex set of probability functions c, defined on $\mathcal{A}$.

EPISTEMIC IMPRECISE DECISION THEORY

An epistemic decision problem (e.d.p.) is characterized by

- a (finite) space W of possible worlds (for the agent),
- an option space $\mathcal{B} : 2^{\mathcal{A}}$ for some agenda $\mathcal{A}$,
- an epistemic utility function $U : \mathcal{B} \times W \mapsto \mathbb{R}$.

Epistemic "choice" is then a choice between all the different combinations of categorical attitudes that you may have towards the propositions of which you are aware.

The utility of a belief set is identified with its degree of accuracy: How well the categorical attitudes it encode reflect the actual world.

$$U(w, B) = \sum_{p \in \mathcal{A}} v(B, p, w)$$

$$v(B,p,w) = \begin{cases} R & \text{if } p \in B \text{ and } p \text{ is \textbf{true} at } w \\ 0 & \text{if } p \notin B \\ W & \text{if } p \in B \text{ and } p \text{ is \textbf{false} at } w \end{cases}$$

where $W, R \in \mathbb{R}$ and $W < 0 < R$.

EU_c gives the expected accuracy of a belief set, given a probability function c:

$$EU_c(B) = \sum_{w \in W} c(\{w\}) U(w, B)$$

IEU_C gives the expected accuracy of a belief set given a credal set C:

$$IEU_C(B) = \{EU_c(B) \mid c \in C\}$$

LOCKEANISM FOR PRECISE PROBLEMS

Maximize expected accuracy = Maximize expected utility, with the assumption that epistemic utility = accuracy.

$$\text{MEA}(\mathcal{B}) = \{B \in \mathcal{B} \mid \forall B' \in \mathcal{B} : EU_c(B) \geq EU_c(B')\}$$

A belief state B will maximize your expected accuracy iff, for all propositions p on your agenda:

- $p \in B$ iff $c(p) \geq \frac{-W}{R-W}$,
- $\neg p \in B$ iff $c(p) \leq 1 - \frac{-W}{R-W}$, and
- $p, \neg p \notin B$ iff $(1 - \frac{-W}{R-W}) \leq c(p) \leq \frac{-W}{R-W}$.

Henceforth: $\frac{W}{R-W}$ = bt (belief threshold), and $1 -$ bt = dc (disbelief ceiling).

References (paper). Dorst, K. (2019). Lockeans maximize expected accuracy. *Mind* 128(509), 175–211. | Easwaran, K. (2016). Dr. Truthlove, or: How I learned to stop worrying and love Bayesian probabilities, *Noûs* 50(4), 816–853. | Foley, R. (1992). The epistemology of belief and the epistemology of degrees of belief, *American Philosophical Quarterly* 29(2), 111–124. | Gilboa, I. and Schmeidler, D. (1989). Maximin expected utility with non-unique prior, *Journal of mathematical economics*. | Levi, I. (1974). On Indeterminate probabilities, Journal of Philosophy, 71: 391–418. | Walley, P. (1991), *Statistical reasoning with imprecise probabilities*, Chapman & Hall.

GENERALIZING LT

We consider three well-known rules for imprecise decision making, which all agree with MEA when expectation values are precise.

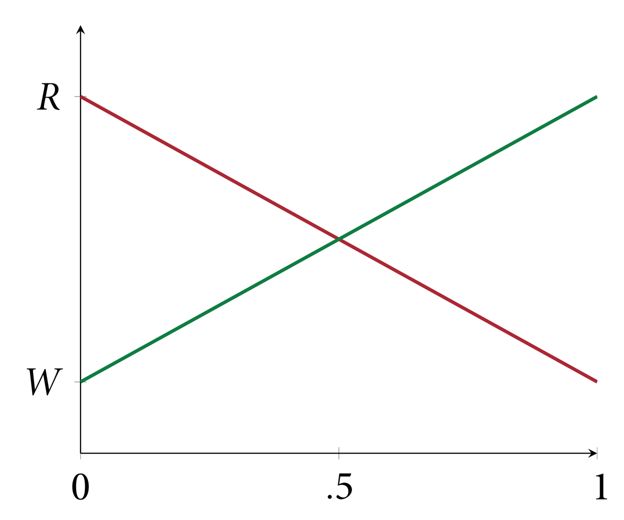
E-admissibility. $\text{E}(\mathcal{B}) = \{B \in \mathcal{B} \mid \exists c \in C, \forall B' \in \mathcal{B} : EU_c(B) \geq EU_c(B')\}$

Γ-maximin. $\text{MAXIMIN}(\mathcal{B}) = \{B \in \mathcal{B} \mid \forall B' \in \mathcal{B} : \underline{IEU}_C(B) \geq \underline{IEU}_C(B')\}$

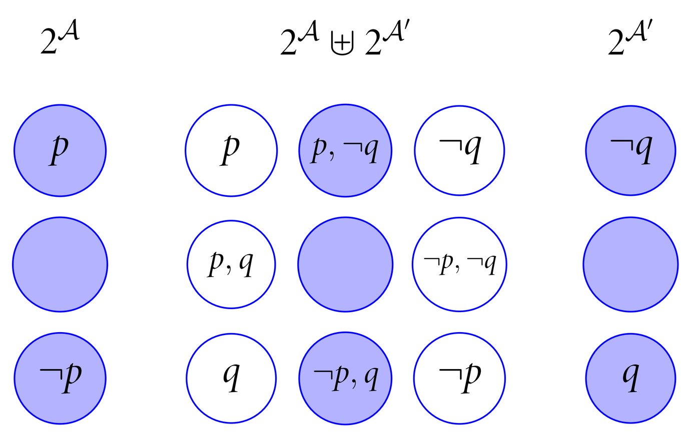
Maximality. $\text{MAXI}(\mathcal{B}) = \{B \in \mathcal{B} \mid \forall B' \in \mathcal{B}, \exists c \in C : EU_c(B) \geq EU_c(B')\}$

Given the minimal agenda $\mathcal{A} = \{p, \neg p\}$:

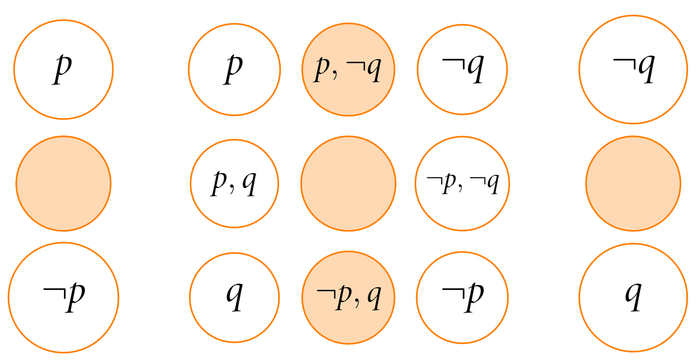
	E-admissibility	Γ-maximin	Maximality
believe	$\overline{C}(p) \geq$ bt	$\underline{C}(p) \geq$ bt	$\overline{C}(p) >$ bt
disbelieve	$\underline{C}(p) \leq$ dc	$\overline{C}(p) \leq$ dc	$\underline{C}(p) <$ dc
suspend	$C(p) \cap [\text{dc}, \text{bt}] \neq \emptyset$	$C(p) \not\subseteq [\text{dc}, \text{bt}]$	$C(p) \cap (\text{dc}, \text{bt}) \neq \emptyset$



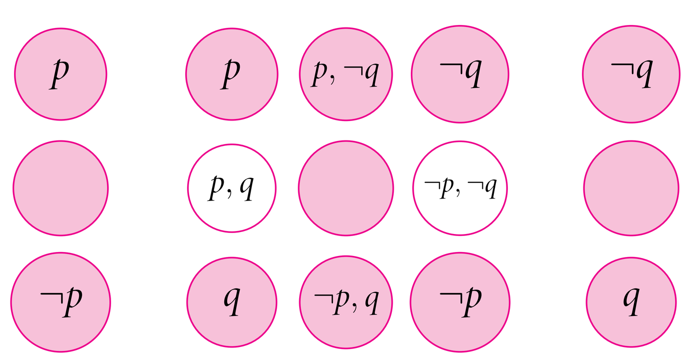
Space to draw! Compare predictions at different credence intervals and suspension plots.



Case 1: E-admissibility. Consider C with limits $c_1, c_2 : c_1(p) = c_1(\neg q) = c_2(q) = c_2(\neg p) > .5$, and U with $v : R + W > S$.



Case 2: Γ-maximin. Just like Case 1 except $v : R + W = S$.



Case 3: Maximality. Just like Case 1.

PERSISTENCE

The Lockean thesis implies that rational categorical attitudes are persistent: If attitude φ towards p is rationally permitted given some agenda, it should remain permitted if we add or remove other propositions (as long as this does not affect our credences).

Define

$$A \uplus B = \{a \cup b \mid a \in A \text{ and } b \in B\}.$$

We use this to define the following properties of a decision rule f. The first two are adaptions of the Sen's property α and β, and are easily seen to together imply Persistence.

C-consistency. $f(\mathcal{B} \uplus \mathcal{B}) \subseteq f(\mathcal{B}) \uplus f(\mathcal{B}')$.

E-consistency. $f(\mathcal{B}) \uplus f(\mathcal{B}') \subseteq f(\mathcal{B} \uplus \mathcal{B}')$

Persistence. $f(\mathcal{B} \uplus \mathcal{B}') = f(\mathcal{B}) \uplus f(\mathcal{B}')$

Neither of our rules are persistent. E and MAXI are C- but not E-consistent, while MAXIMIN is E- but not C-consistent.

Relevant counterexamples are illustrated to the left. Option spaces are built from the agendas $\mathcal{A} = \{p, \neg p\}$, $\mathcal{A}' = \{q, \neg q\}$, and their union (inconsistent options removed for simplicity). Circles depict the available belief sets for the given option space, and shading marks permission by the decision rule in question.

DO WE WANT PERSISTENCE?

Claim: When all credences are precise, persistence is appealing for both synchronic reasons (enabling reductions, ensuring coherent demands on action) and diachronic reasons (rational awareness growth/reduction). But it is not generally desirable once we introduce imprecision:

In Case 1 and 3, you are certain that p and q are unlikely to share truth value. In those examples, you dislike believing falsehoods more than you like believing truths, making $\{p, q\}, \{\neg p, \neg q\}$ intuitively irrational options.

In Case 2, you are as unsure of p as you are of its negation. Unless you value true beliefs more than you disvalue false ones, suspending on p is intuitively rational.

Rejecting persistence does not preclude agenda-independent rationality constraints, since persistence violations can still ultimately be traced to properties of our graded beliefs. Broadly, they occur with credal sets C such that for ≤ 2 propositions, the members of C

- disagree on which individual proposition is the likelier to be true, but
- agree that the propositions are unlikely to share truth value.

Going forward, a main goal will be to find a stringent characterization of these credal sets, in order to define a notion of partial persistence suitable for the imprecise case.