THE JEFFREY-BOLKER DECISION FRAMEWORK

DANIEL A. HERRMANN & GERARD ROTHFUS

Overview

The goal of this talk is the introduce to Jeffrey-Bolker decision theory framework, and show that it has some useful features. We start by introducing Savage's decision theory framework, and discussing some of its shortcomings. We then introduce the Jeffrey-Bolker framework and describe how it overcomes the problems of Savage's framework. We highlight the weaker uniqueness result of Bolker's representation theorem, and describe ways that it can be strengthened. We end with highlighting two aspects of the Jeffrey-Bolker system. The first is its connection to both formal and conceptual developments in measure theory; the second is its ability to accommodate different decision theories (evidential, causal, etc.).

1. The Savage Framework

Savage envisions decision problems as comprised of three main components:

- A set of *states of the world*, S
- A set of *consequences*, C
- A set of *acts*, $X \subseteq A = C^S$

An agent's preferences are assumed to extend across all hypothetical acts (not merely those in her immediate choice set X) and are captured in a binary *preference relation*, $\succeq$, on A, with $f \succeq g$ indicating that the agent does *not* strictly prefer g to f. The symmetric and asymmetric parts of this relation are denoted by $\sim$ and $\succ$, respectively. Note: we can extend $\succeq$ to C by identifying consequences with the constant acts that yield them.

P 1 (Ordering)**.** *$\succeq$ is a weak order: it is transitive and complete.*

P 2 (The Sure-Thing Principle)**.** *For all $f, g, h, h' \in A$ and $E \subseteq S$,*
$f^h_E \succeq g^h_E$ if and only if $f^{h'}_E \succeq g^{h'}_E$.[1]

P 3 (State Neutrality)**.** *For all $f \in A$, nonnull events $E \subseteq S$, and $x, y \in C$,*
$x \succeq y$ if and only if $f^x_E \succeq f^y_E$.[2]

[1] Notation: If $f, h \in A$ and $E \subseteq S$, let f^h_E be the act that agrees with f on E and with h outside of E.
[2] An event E is *null* just in case indifference holds between any two acts that agree outside of E.

P 4. *For all* $E_1, E_2 \in S$, *and* $x, y, z, w \in C$ *with* $x \succ y$ *and* $z \succ w$,
$y^x_{E_1} \succeq y^x_{E_2}$ *if and only if* $w^z_{E_1} \succeq w^z_{E_2}$.

P 5 (Non-triviality)**.** *There exist* $f, g \in A$ *such that* $f \succ g$.

P 6 (Non-atomicity)**.** *For all* $f, g, h \in A$ *with* $f \succ g$, *there exists a partition of* S, $\{E_1, ..., E_n\}$, *such that for every* $i \leq n$, $f^h_{E_i} \succeq g$ *and* $f \succeq g^h_{E_i}$.

P 7. *For all* $f, g \in A$ *and events* $E \in S$, *(i) if for every* $s \in E$, $f \succeq_E g(s)$, *then* $f \succeq_E g$, *and (ii) if for every* $s \in E$, $g(s) \succeq_E f$, *then* $g \succeq_E f$.[3]

Theorem 1 (Savage's Existence Theorem)**.** *P1-P7 hold if and only if there exists a nonatomic*[4] *finitely-additive probability measure* P *on* $\mathcal{S}$ *and a nonconstant bounded utility function* $u : C \to \mathbb{R}$ *such that for every* $f, g \in A$,
$f \succeq g$ *if and only if* $\int_S u(f(s))dP(s) \geq \int_S u(g(s))dP(s)$

This entails that for any two *simple* acts, f and g, that yield constant output within each cell of some finite partition, $\{E_1, ..., E_n\}$,
$f \succeq g$ if and only if $\sum_i u(x_i)P(E_i) \geq \sum_i u(y_i)P(E_i)$,
where x_i and y_i are are the consequences yielded by f and g, respectively, within event E_i.

Theorem 2 (Savage's Uniqueness Theorem)**.** P *is unique and* u *is unique up to positive affine transformation. That is, if* (P, u) *represents* $\succeq$ *in the manner described above, then* (P', u') *does so as well if and only if* $P = P'$ *and* $u = au' + b$ *for some* $a \in \mathbb{R}$ *and* $b > 0$.

2. Problems with the Savage Framework

Despite its many virtues, Savage's framework encounters some objections.

(1) *Alleged Counterexamples*: Allais, Ellsberg, etc.
(2) *State-Consequence Independence*: P2 and P3 are only plausible if consequences are characterized to a sufficient level of detail to ensure their desirabilisitic independence from the state in which they are realized.
(3) *Rectangular Field Assumption I*: The agent has a preference ranking over every possible function from states to consequences. However, given the requirement of state-consequence independence, such functions may well include the logically bizarre and difficult to interpret.
(4) *Rectangular Field Assumption II*: Relatedly, the fact that an agent must be able to express preferences regarding the vast range of acts Savage

[3] Notation: $f \succeq_E g$ just in case $f' \succeq g'$ where f' and g' agree, respectively, with f and g on E but agree with one another off E. Satisfaction of the STP guarantees that choice of such f' and g' doesn't matter and so $\succeq_E$ is well-defined.

[4] A probability measure P is *nonatomic* just in case every event E it is defined on is such that $P(E) > 0$ implies the existence of an $F \subset E$ such that $P(E) \neq P(F)$ and $P(F) > 0$.

requires may require her, in making such judgments, to revise her beliefs regarding what is causally possible in radical ways.

(5) *Dualism*: Savage draws a sharp distinction between states, consequences, and acts; of these only states are assigned probabilities. The agent does not assign probabilities to her own acts; this is a clear way in which she does not model herself as part of the world.

(6) *Act-State Independence*: Savage's framework seeks to measure an agent's epistemic attitudes; but to get off the ground it needs to set up the acts and states such that the agent views them as independent. It thus can't be used to measure this independence.

(7) *Formulation Invariance*: The above requirements (like state-consequence desirabilisitic independence and act-state probabilistic independence) render the applicability of Savage's decision theory highly sensitive to how states and consequences are formulated.

3. The Jeffrey-Bolker Framework

One core innovation of the Jeffrey-Bolker framework is that everything—acts, states, consequences—is the same type of object. This ends up erasing many distinctions found in other theories, such as Savage's. Thus, the preference ranking is defined over one single algebra.

Definition 1. *If $\mathfrak{A}$ is an algebra, define $\mathfrak{A}'$ as $\mathfrak{A} - F$, where F is the bottom element of $\mathfrak{A}$. A Boolean algebra $\mathfrak{A}$ is **atomless** iff $\forall A \in \mathfrak{A}', \exists B \in \mathfrak{A}' : B \to A$. A Boolean algebra is **complete** iff every subset of the algebra has both a supremum and infimum in the algebra.*

Axiom 1. *$\succeq$ is defined on $\mathfrak{A}'$, where $\mathfrak{A}$ is a complete, atomless Boolean algebra.*

Remark 2. *We interpret $A \succeq B$ as saying "B is not preferred to A" or equivalently, "A is at least as preferred as B". $\succ$ and $\sim$ are defined in the obvious ways.*

Axiom 3. *$\succeq$ is a complete preorder: it is fully connected, reflexive, and transitive.*

Axiom 4 (Averaging)**.** *If $A, B \in \mathfrak{A}'$ and $A \wedge B = F$, then:*

(1) *if $A \succ B$, then $A \succ A \vee B \succ B$; and*
(2) *if $A \sim B$, then $A \sim A \vee B \sim B$.*

Axiom 5 (Impartiality)**.** *For all $A, B \in \mathfrak{A}'$, whenever $A \wedge B = F$ and $A \sim B$, then, if $A \vee C \sim B \vee C$ for some $C \in \mathfrak{A}'$ such that $A \wedge C = B \wedge C = F$, and $C \not\sim A$, then $A \vee D \sim B \vee D$ for every $D \in \mathfrak{A}'$ disjoint from A and B.*

Averaging ensures that the disjunction of two propositions lies between the two propositions.[5] Impartiality basically gives a way to test that the agent views two disjoint propositions (A and B) as equally probable.[6]

Axiom 6 (Continuity)**.** *Let $\mathcal{A} = \{A_1, A_2, \ldots\} \subseteq \mathfrak{A}$ be a sequence of propositions such that $A_n \to A_{n+1}$, for all n. Then, if A^* is the supremum of $\mathcal{A}$ and $B \succ A^* \succ C$, then there is some N such that $B \succ A_n \succ C, \forall n \geq N$. The analogous condition holds for decreasing sequences of propositions.*

We call a preference ordering that satisfies all of the axioms *coherent.*

Theorem 7 (Bolker's Existence Theorem.)**.** *Let $\mathfrak{A}$ be a complete atomless Boolean algebra, and let $\succeq$ be a coherent preference ordering on $\mathfrak{A}'$. Then there exists a probability measure P defined on $\mathfrak{A}$ and a signed measure v on $\mathfrak{A}$ such that, for all A and B in $\mathfrak{A}'$:*

$$A \succeq B \text{ iff } U(A) \geq U(B)$$

where

$$U(A) = \frac{v(A)}{P(A)}, \forall A \in \mathfrak{A}'.$$

It follows that, for any $A \in \mathfrak{A}'$ and partition $\mathcal{S} = \{S_i\}_{i \in I} \subset \mathfrak{A}$ of A, we have that we can calculate the utility of A as follows:

$$U(A) = \sum_{i \in I} U(A \wedge S_i) P(S_i | A).$$

This makes clear that the representation is in fact one of expected utility.

Theorem 8 (Bolker's Uniqueness Theorem.)**.** *Let P, P' be probability measures on $\mathfrak{A}$ and let v, v' be signed measures on $\mathfrak{A}$. Then the pair P', v' represents the same preference order as P, v iff:*

$$v' = av + bP, \text{ and } P' = cv + dP$$

where $ad - bc > 0$, $cv(T) + d = 1$, and for all $A \in \mathfrak{A}'$, $-\frac{d}{c} \neq \frac{v(A)}{P(A)}$.

This transformation of P and v to P' and v' induces the following shift in U:

---

[5] If you squint hard enough, then Averaging is somewhat similar to the irrelevance of independent alternatives axioms present in other utility theories, and Savage's Sure-Thing Principle. But they are also very different.

[6] Suppose A and B are disjoint (meaning $A \wedge B = F$), and suppose that the agent is indifferent between them ($A \sim B$). In the background, we are imagining that the preference ranking has arisen by the principle of maximizing expected utility. Then, the agent will be indifferent between $A \vee C$ and $B \vee C$ only in the case where A and B are equiprobable. For, suppose not. For concreteness, imagine that $C \succ A$, and that B is more probable than A. Then the agent will prefer $A \vee C$ to $B \vee C$, since this gives the agent a higher probability of getting the more desired outcome, C. This is how this condition works as a kind of test for equiprobability. The axiom says that this test will come out the same way, *no matter which* test proposition D is used.

$$U' = \frac{v'}{P'} = \frac{av + bP}{cv + dP} = \frac{aU + b}{cU + d}.$$

4. How Jeffrey-Bolker Addresses the Savage Limitations

But why not be a bit more holistic, and view the agent as part of nature. A state of nature would then specify what act the agent performs, along with everything else one usually takes it to specify.
∼ Richard Jeffrey, *Frameworks for preferences*

We can see that the Jeffrey-Bolker Framework addresses some of the previously mentioned issues with the Savage framework.

(1) *State-Consequence Independence*: This is not required.
(2) *Dualism*: The agent assigns probabilities to propositions about her own acts like anything else.
(3) *Rectangular Field Assumptions I & II*: The agent's preferences are defined on a set that is closed only on *logical* operations, not causal (or functional) operations. This allows the preferences to only be defined on propositions that the agent takes to possible.
(4) *Act-State Independence*: This is not required.
(5) *Formulation Invariance*: Jeffrey-Bolker is partition invariant, meaning you can calculate with respect to any partition and get the same answer.

5. Strengthening the Uniqueness Result

There are two main approaches to strengthening Bolker's Uniqueness Theorem to more closely resemble Savage's.

(1) *Unbounded Utility*: If desirabilities are unbounded, that is, if the range of U is unbounded above and below, we can secure the uniqueness of P. Jeffrey offers the following as a qualitative condition on preference that is, in the presence of the preceding axioms, necessary and sufficient for U to be unbounded in the manner sufficient for P's uniqueness.
   - The Unboundedness Condition: Indifference holds between all propositions D that satisfy conditions a and b below are ranked together.
     (a) Whener A, B, C are pairwise incompatible propositions of which A and B are ranked with T and C is ranked above $A \vee C$, which in turn is ranked with $B \vee C$ and with or above G, the proposition $A \vee B \vee C$ is ranked above D.[7]
     (b) Whener A, B, C are pairwise incompatible propositions of which A and B are ranked with G and C is ranked below $A \vee C$, which

[7] G is an arbitrarily chosen proposition such that $G \succ T \succ \neg G$.

in turn is ranked with $B \vee C$ and with or below T, the proposition $A \vee B \vee C$ is ranked below D.

(2) *Comparative Probability*: Another approach to securing the uniqueness of P involves introducing some sort of primitive comparative probability relation into our modelling framework.
   - Ahmed's Version: Introduce an equivalence relation, $\approx$, on $\mathfrak{A}$. $A \approx B$ is interpreted to mean that the agent is equally confident of A and B. If we (i) require that $A \approx B$ for some $A, B \in \mathfrak{A}$ such that $V(A) > V(B)$, and (ii) assume that among candidate pairs (P, U) that represent $\succeq$ at least one is such that $P(X) = P(Y)$ whenever $X \approx Y$, then we can secure the result that there is a unique probability measure P and unique up to positive affine transformation utility function U that both represents $\succeq$ in the required expectational fashion and agrees with $\approx$.

(3) *Conditionals*: A final approach, proposed by Richard Bradley, involves enriching Jeffrey-Bolker's Boolean algebra to include indicative conditionals and then supplying axiomatic constraints on agent's preferences over such conditionals.

6. Connection to Measure Theory

Here we aim to draw attention to the use of a complete atomless Boolean algebra. While this may seem like a very specific and odd choice, driven by the the representation theorem, there are in fact many reasons to favour such a choice. This is because there is a deep connection here to *measure algebras*.

Definition 2. *A* ***measure algebra*** *is a pair* $(\mathfrak{B}, \mu)$ *where* $\mathfrak{B}$ *is a Boolean* σ*-algebra and* μ *is a strictly positive probability measure on* $\mathfrak{B}$*.*

A common way to construct a measure algebra is to start off with the Lebesgue measurable sets on the unit interval, and then quotient out by the measure zero σ-ideal. Thinking of this procedure, one can see that this procedure destroys all fine-grained measure zero structure, and that the resulting structure does not contain any singletons of possible worlds (which would be atoms).

Measure algebras constructed in this way are *complete* and *atomless*, and so are σ-isomorphic to $\mathfrak{A}$ used in the Jeffrey-Bolker framework. Thus, we might ask, are there independent motivations for using measure algebras for probability spaces?

Kolmogorov, *contra* his original foundational framework for probability theory using what he calls the "set theoretical" system of probability, advocated for measure algebras. Two advantages he points out are of interest for us. The advantages are understood as fixing two defects of the set-theoretical system.

> *1*st, the notion of an elementary event is an artificial superstructure imposed on the concrete notion of an event. In reality, events are not composed of elementary events, but elementary events originate in the dismemberment of composite events. . . .
> *3*rd, we are forced to give up the principle, formulated in numerous classical works in probability theory, according to which an event of probability zero is absolutely impossible.. . .
> All of the inconveniences can be avoided if we base probability theory on [measure algebras]. (p.61, 148, *Algèbres de Boole métriques complètes*, translated by Richard Jeffrey in 1995 for *Philosophical Studies*)

7. Modeling Many Decision Theories

The Jeffrey-Bolker framework, as we've presented it, is simply a framework for modelling rational attitudes, specifically preferences and (if we wish) comparative probabilities. But it is most commonly associated with a particular account of rational choice as well, namely, *Evidential Decision Theory (EDT).* This is the most natural decision rule one can state within the Jeffrey-Booker setup, and the one favored by Jeffrey.

Given a decision problem characterized by a partition of act propositions, choose one that maximizes U. But those sympathetic to the Jeffrey-Booker framework are free to consider other decision rules as well. Newcomb's Problem is the standard motivation for doing so.

Philosophers like Joyce and Bradley have sought to extend the Jeffrey-Bolker system to incorporate distinctively causal elements with an aim toward stating a causal decision rule. Joyce invites us to think of the instrumental value of an act as corresponding to the utility of the tautology on the supposition of its performance. This is in line with EDT, if we understand supposition in terms of learning and define that $U(A|B) = U(A\&B)$:

$EDT(A) = U(A) = U(T\&A) = U(T|A)$

But, Joyce suggests, the correct form of supposition to employ here comes apart from learning. For causal decision theorists, the requisite sort of supposition is *subjunctive* or *interventional.* To find the instrumental value of A, I assess the utility of the status quo that would result were I (perhaps contrary to fact) to intervene to make A true.

Formally, Joyce imagines that when faced with a decision problem given by a partition of acts, agents have suppositional preferences and comparative probabilities given by a family of preference relation/comparative probability pairs

$(\succeq^X, \trianglerighteq^X)$ that each satisfy the Jeffrey-Booker axioms and are so each represented by a pair (U^A, P^A). P^A is often called an *imaging function* and captures the probability of various propositions on the subjunctive supposition that A. Rational choice then goes by:

$CDT(A) = U^A(T) = \sum_S P^A(S|T)U^A(S\&T) = \sum_S P^A(S)U^A(S)$

One standard way to derive P^A is to employ a *K-partition*, a privileged partition of propositions that are assumed to be causally independent of each act A and to determine the causal impact of each act A. We then define:

$P^A(B) = \sum_i P(B|A\&K_i)P(K_i)$

$U^A(B) = \sum_i U(B|A\&K_i)P(K_i)$

These can then determine CDT values and suppositional preferences using U and P.

Joyce's Existence and Uniqueness Theorems show that agents whose suppositional preferences satisfy various axioms can be representable as CDT agents. This set up is significantly more complicated and demanding than EDT, which, beyond the resources the Jeffrey-Booker setup itself provides, only requires a specification of an act partition. However, it continues to enjoy many advantages provided by the Jeffrey-Bolker framework (act probabilities are allowed, no awkward rectangular field assumptions, etc.)