Preference Based Belief Dynamics

Eric Pacuit

ILLC, University of Amsterdam

staff.science.uva.nl/∼epacuit

epacuit@science.uva.nl

April 27, 2006

Introduction

- Starting with the work of Savage (based on Ramsey and de Finetti), there is a tradition in game theory and decision theory to *define* beliefs and utilities in terms of the agent's preferences

- Thus logical properties of beliefs are derived from properties of the preferences.

- Typically the results come in the form of a representation theorem:

  > If the agents preferences satisfy such-and-such properties, then there is a set of conditional probability functions and a (state independent) utility function such that the agent can be assumed to act as an *expected utility maximizer*.

Outline of the Talk

- Sketch Savage's Theorem
- Defining Beliefs from Preferences (S. Morris)
- Deducing Logical Properties of Beliefs from Properties of Preferences (S. Morris)
- Pointers to future work (with Olivier Roy)

Expected Utility Theorem: Notation

- Let A be an arbitrary set, then $\Delta(A)$ is the set of a probability distributions on A.

- Let X be a *finite* set of **prizes** or **outcomes**.

- Let Ω be a *finite* set of **states**

- A **lottery** is a funciton $f : \Omega \to \Delta(X)$.

  The intended interpretation of $f(x|t)$ is "the *objective* probability of getting prize x given that the current state is t".

- For each $x \in X$, define $[x](y|t) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$

Expected Utility Theorem: Notation

- Let $\Sigma = \{S \mid S \subseteq \Omega \;\&\; S \neq \emptyset\}$ is a set of events.

- Given a set $S \in \Sigma$ and lotteries f and g, $f \preceq_S g$ is intended to mean that "g is at least as good as f, given that the true state of the world is in S."

  If the agent thinks the actual state is in S, then the agent would choose lottery f over g.

- $f \sim_S g$ and $f \prec_S g$ are defined as usual.

Expected Utility Theorem: Notation

- A **conditional-probability function** $p : \Sigma \to \Delta(\Omega)$ is a function that gives the probability of a state t given that an event $S \in \Sigma$ occurs.

- A **utility function** is any function $u : X \times \Omega \to \mathbb{R}$. It is said to be *state independent* iff for all $s, t \in \Omega, u(x, s) = u(x, t)$.

- The **expected utility value** of a lottery f given an even S, $E_p(u(f)|S)$, is calculated as

$$E_p(u(f)|S) = \sum_{t \in S} p(t|S) \sum_{x \in X} u(x, t) f(x|t)$$

Expected Utility Theorem

Let f, g be lotteries and S and event.

1. *(Completeness)* $f \succeq_S g$ or $g \succeq_S f$.

2. *(Transitivity)* If $f \succeq_S g$ and $g \succeq_S h$ then $f \succeq_S h$.

3. *(Relevance)* If, for all $t \in \Omega$, $f(\cdot|t) = g(\cdot|t)$ then $f \sim_S g$.

4. *(Monotonicity)* If $f \succ_S g$ and $0 \leq \beta < \alpha \leq 1$ then $\alpha f + (1-\alpha)g \succ_S \beta f + (1-\beta)g$.

5. *(Continuity)* If $f \succeq_S g$ and $g \succeq_S h$ then there exists a number γ such that $0 \leq \gamma \leq 1$ and $g \sim_S \gamma f + (1-\gamma)h$.

Expected Utility Theorem

6. *(Objective substitution)* If $e \succeq_S f$, $h \succeq_S g$ and $0 \leq \alpha \leq 1$, then $\alpha e + (1-\alpha)h \succeq_S \alpha f + (1-\alpha)g$.

7. *(Subjective substitution)* If $g \succeq_S f$, $g \succeq_T f$ and $S \cap T = \emptyset$ then $g \succeq_{S \cup T} f$.

8. *(Non-triviality)* For all $t \in \Omega$ there exist $x, y \in X$ such that $[x] \succ_{\{t\}} [y]$

9. *(State Neutrality)* For all $r, t \in \Omega$, if $f(\cdot|r) = f(\cdot|t)$, $g(\cdot|r) = g(\cdot|t)$ and $g \succeq_{\{r\}} f$, then $g \succeq_{\{t\}} f$.

Expected Utility Theorem

6. *(Objective substitution)* If $e \succ_S f$, $h \succeq_S g$ and $0 < \alpha \leq 1$, then $\alpha e + (1-\alpha)h \succ_S \alpha f + (1-\alpha)g$.

7. *(Subjective substitution)* If $g \succ_S f$, $g \succ_T f$ and $S \cap T = \emptyset$ then $g \succ_{S \cup T} f$.

8. *(Non-triviality)* For all $t \in \Omega$ there exist $x, y \in X$ such that $[x] \succ_{\{t\}} [y]$

9. *(State Neutrality)* For all $r, t \in \Omega$, if $f(\cdot|r) = f(\cdot|t)$, $g(\cdot|r) = g(\cdot|t)$ and $g \succeq_{\{r\}} f$, then $g \succeq_{\{t\}} f$.

Expected Utility Theorem

6. *(Objective substitution)* If $e \succeq_S f$, $h \succeq_S g$ and $0 \leq \alpha \leq 1$, then $\alpha e + (1-\alpha)h \succeq_S \alpha f + (1-\alpha)g$.

7. *(Subjective substitution)* If $g \succeq_S f$, $g \succeq_T f$ and $S \cap T = \emptyset$ then $g \succeq_{S \cup T} f$.

8. *(Non-triviality)* For all $t \in \Omega$ there exist $x, y \in X$ such that $[x] \succ_{\{t\}} [y]$

9. *(State Neutrality)* For all $r, t \in \Omega$, if $f(\cdot|r) = f(\cdot|t)$, $g(\cdot|r) = g(\cdot|t)$ and $g \succeq_{\{r\}} f$, then $g \succeq_{\{t\}} f$.

Expected Utility Theorem

Axioms 1 - 8 are jointly satisfied iff there exists a utlity function $u : X \times \Omega \to \mathbb{R}$ and a conditional probability function $p : \Sigma \to \Delta(\Omega)$ such that

1. For all $t \in \Omega$, $\max_{x \in X} u(x,t) = 1$ and $\min_{x \in X} u(x,t) = 0$.
2. For all R, S, T such that $R \subseteq S \subseteq T \subseteq \Omega$ and $S \neq \emptyset$, $p(R|T) = p(R|S)p(S|T)$.
3. For all $f, g \in L$ and $S \in \Sigma$, $f \succeq_S g$ iff $E_p(u(f)|S) \geq E_p(u(g)|S)$.

If, furthermore, axiom 8 is also satisfied, then u is state-independent.

Expected Utility Theorem

Axioms 1 - 8 are jointly satisfied iff there exists a utlity function $u : X \times \Omega \to \mathbb{R}$ and a conditional probability function $p : \Sigma \to \Delta(\Omega)$ such that

1. For all $t \in \Omega$, $\max_{x \in X} u(x,t) = 1$ and $\min_{x \in X} u(x,t) = 0$.
2. For all R, S, T such that $R \subseteq S \subseteq T \subseteq \Omega$ and $S \neq \emptyset$, $p(R|T) = p(R|S)p(S|T)$.
3. For all $f, g \in L$ and $S \in \Sigma$, $f \succeq_S g$ iff $E_p(u(f)|S) \geq E_p(u(g)|S)$.

If, furthermore, axiom 8 is also satisfied, then u is state-independent.

There is a large literature surrounding this result!

For this Talk

Let Ω be a set of states.

Let X be a finite set of prizes.

An **act** is a function $x : \Omega \to \mathbb{R}$. Let $\mathbb{R}^{\Omega}$ be the set of all acts.

x_w for $w \in \Omega$ means that if the true state is w, then the agent receives prize x.

We write $x \succeq_w y$ the agent prefers x over y *provided the true state is* w

For this Talk

Let Ω be a set of states.

Let X be a finite set of prizes.

An **act** is a function $x : \Omega \to \mathbb{R}$. Let $\mathbb{R}^\Omega$ be the set of all acts.

x_w for $w \in \Omega$ means that if the true state is w, then the agent receives prize x.

We write $x \succeq_w y$ the agent prefers x over y *provided the true state is w*

Belief Operators

A **belief operator** is a function $B : 2^\Omega \to 2^\Omega$

For $E \subseteq \Omega$, $w \in B(E)$ means the agent believes E at state w

B is normal if

- $B(\Omega) = \Omega$
- $B(E \cap F) = B(E) \cap B(F)$

Possibility function: $P : \Omega \to 2^\Omega$: set of states the agent considers possible at w

$\mathsf{b} : \Omega \to 2^{2^\Omega}$ lists the agents belief state at w.

Defining Beliefs from Preferences

For $E \subseteq \Omega$ and two acts x and y, let (x_E, y_{-E}) denote the new act that is x on E and y on $-E$.

B **reflects** $\{\succeq_w\}_{w \in \Omega}$ provided for each $E \subseteq \Omega$

$$B(E) = \{w \mid (x_E, y_{-E}) \sim_w (x_E, z_{-E}) \quad \text{for all } x, y, z \in \mathbb{R}^{\Omega}\}$$

Theorem If the preference relations are complete and transitive, then the derived belief operator is normal.

If the preferences have a state independent expected utility representation, then

$$B(E) = \{w \mid \sum_{w' \in E} p(w'|w) = 1\}$$

$$P(w) = \{w' \mid p(w'|w) > 0\}$$

Some Notation

For $x, y \in \mathbb{R}^\Omega$,

$x \geq y$ iff for each $w \in \Omega$, $x_w \geq y_w$

$x > y$ iff $x_w \geq y_w$ for each $w \in \Omega$ and $x_{w'} > y_{w'}$ for some $w' \in \Omega$

$x >> y$ iff $x_w > y_w$ for each $w \in \Omega$

Alternative Definitions

$$B^*(E) = \{w \mid (x_E, z_{-E}) \succeq_w (y_E, v_{-E}) \text{ for all } x >> y,\ x, y, z, v \in \mathbb{R}^\Omega\}$$

Preferences are **monotone** if $x >> y$ implies $x \succ_w y$ and $x \geq y$ implies $x \succeq_w y$ for all $w \in \Omega$.

Theorem B^* is normal if the preference relations are monotone, non-trivial and transitive.

$$P^*(w) = \{w' \mid (x_{w'}, z_{-w'}) \succeq_w y \text{ for some } x >> y >> z\}$$

Lexiographic Expected Utility Representation

Take into account probability zero events.

Preferences relation $\{\succeq_w\}_{w\in\Omega}$ have a LEU representation if there is a positive integer J such that for each $w \in \Omega$ and $j = 1, \ldots, J$, probability distributions $p_j(\cdot \mid w)$ on Ω such that

$$x \succeq_w y \Leftrightarrow \left\{ \sum_{w'\in\Omega} p_j(w'|w) u_w(x'_w) \right\}_{j=1}^{J} \geq_L \left\{ \sum_{w'\in\Omega} p_j(w'|w) u_w(y'_w) \right\}_{j=1}^{J}$$

B is belief with probability 1 for each $j = 1, \ldots, J$

B^* is probability 1 for p_1

See S. Morris, Alternative Definitions of Knowledge in *Epistemic Logic and the Foundations of Decision and Game Theory*, eds. Bacharach et al.

Additional Properties of Preferences

Continuity The set $\{x \in \mathbb{R}^\Omega \mid x \succeq y\}$ is closed for all $y \in \mathbb{R}^\Omega$

Monotonicity:

1. If $x \geq y$ then $x \succeq y$
2. If $(x_w, z_{-w}) \succeq (y_w, z_{-w})$ for some $x, y, z \in \mathbb{R}^\Omega$ then $(x_w, z_{-w}) \succ (y_w, z_{-w})$ for all $x, y, z \in \mathbb{R}^\Omega$ such that $x_w > y_w$
3. If $x >> y$ then $x \succ y$

Under these conditions,

$$P(w) = \{w' \mid \text{for all } x >> y \text{ there exists } z << y \text{ such that } (x_{w'}, z_{-w'}) \succ_w y\}$$

Coherency

A minimal rationality property relating together preferences at different states of Ω.

Preferences are **coherent** if choices made at different states can be seen as reflecting a true, metapreference ordering over acts.

A preference relation $\succ$ is a **meta-ordering** if it is complete, transitive, continuous and for all $x, y, z \in \mathbb{R}^{\Omega}$

$$(x_w, z_{-w}) \succeq (y_w, z_{-w}) \Leftrightarrow x_w \geq y_w$$

Decision Problem

A decision problem is a finite set of acts D

$C_w[D] = \{x \in D \mid x \succeq_w y \text{for all } \; y \in D\}$

A decision rule is a function $f : \Omega \to D$

f is optimal provided for each $w \in \Omega$, $f(w) \in C_w[D]$

$C^*[D] = \{x \in \mathbb{R}^\Omega \mid x_w = f_w(w) \text{ for all } w \in \Omega \text{ for some optimal } f\}$

Coherency

Preferences are coherent if there exists a meta-ordering $\succeq_*$ such that for each finite D, there exists $x \in C^*[D]$ such that $x \succeq_* y$, for all $y \in D$.

Theorem If preferences are coherent, then the beliefs reflecting them satisfy the knowledge axiom ($B(E) \subseteq E$) and positive introspection ($B(E) \subseteq B(B(E))$).

Example

Let $\Omega = \{a, b, c\}$, $P(a) = \{a, b\}$, $P(b) = \{b\}$, $P(c) = \{b, c\}$

$x \succeq_a y \Leftrightarrow \frac{1}{2}x_a + \frac{1}{2}x_b \geq \frac{1}{2}y_a + \frac{1}{2}y_b$

$x \succeq_b y \Leftrightarrow x_b \geq y_b$

$x \succeq_c y \Leftrightarrow \frac{1}{2}x_b + \frac{1}{2}x_c \geq \frac{1}{2}y_b + \frac{1}{2}y_c$

$x \succeq_* y \Leftrightarrow \frac{1}{6}x_a + \frac{2}{3}x_b + \frac{1}{6}x_c \geq \frac{1}{6}y_a + \frac{2}{3}y_b + \frac{1}{6}y_c$ is coherent

Example 2

Suppose $\Omega = \{a, b\}$

$x \succeq_a y \Leftrightarrow x_a \geq y_a$

$x \succeq_b y \Leftrightarrow x_a \geq y_a$

$D = \{\mathbf{0}, x\}$, where $x = (\epsilon, -1)$

For ϵ sufficiently small $\mathbf{0} \succeq_* x$, but $C^*[D] = \{x\}$

Example 3

Suppose $\Omega = \{a, b, c\}$

$x \succeq_a y \Leftrightarrow \alpha x_a + (1-\alpha)x_b \geq \alpha y_a + (1-\alpha)y_b$

$x \succeq_b y \Leftrightarrow x_b \geq y_b$

$x \succeq_c y \Leftrightarrow \beta x_a + (1-\beta)x_c \geq \beta y_a + (1-\beta)y_c$

$D = \{\mathbf{0}, x\}$ where $x = (\epsilon, -1, -\epsilon^2)$

$C^*[D] = \{(0, 0, -\epsilon^2)\}$, but $\mathbf{0} \geq_* (0, 0, -\epsilon^2)$ for any meta-ordering.

Dynamic Preferences and Beliefs (Sketch)

Extend preferences to finite time periods $t = 0, \ldots, T$.

Assume preferences for each time period, denoted $\succeq_{w,t}$.

$C_t^*[D] = \{x \in \mathbb{R}^\Omega \mid x_w = f_w(w)$ for some f at time $t\}$

Valuable Information: there is a meta-ordering $\succeq_*$ such that for each finite $D \subseteq \mathbb{R}^\Omega$ and each $t = 1, \ldots, T$, there is a $x(t) \in C_t^*[D]$ such that

$$x(t) \succeq_* y \text{ for all } \; y \in D \cap C_0^*[D] \cap \cdots \cap C_{t-1}^*[D]$$

Dynamic Properties of Beliefs

Historical Negative Introspection
$-B_s(E) \cap B_t(F) \cap -B_s(F) \subseteq B_t(-B_s(E))$ for all events E and F and times $s \leq t$.

Refinement For all $s \leq t$, $B_s(E) \subseteq B_t(E)$.

Theorem If preferences satisfy valuable information, then the induced beliefs satisfy the knowledge axiom, positive introspection, refinement and historical negative introspection.

See S. Morris, The Logic of Belief and Belief Change: A Decision Theoretic Approach, *Journal of Economic Theory*, 1996.

Future and Related Work

1. Adding dyanmics: Different decision problems; same decision problem but new information is received; time stamps vs. movement in the set of states.

2. Properties of preferences change in the face of new information can be lifted to properties of belief change (cf. Hansson, van Benthem and Liu).

3. How does AGM, updates, etc. fit into the above picture? (cf. Perea, Schullte, Ansheim, etc.)

4. Related approach of B. Lipman: Decision theory without logical omniscience.

5. There is a nice survey by Ansheim and Sovik that focuses probabilistic beliefs.

Happy Birthday Prof. Segerberg!