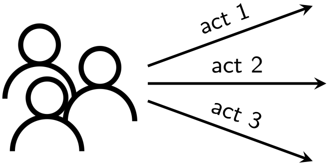
Belief-Averaging and Relative Utilitarianism

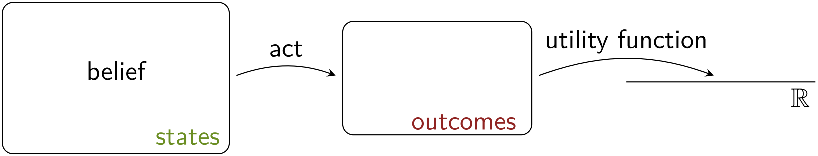
Florian Brandl
University of Bonn

Aggregating preferences under uncertainty



agents' preferences over acts $\longrightarrow$ collective preferences over acts

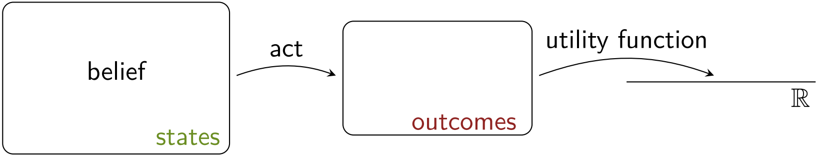
Primitives: states, outcomes, agents



SEU preference: ranking of acts by expected utility for some belief and utility function

Preference profile: tuple of SEU preferences for subset of agents

Aggregation function: function from profiles to SEU preferences



states

		Gentleman 1 is stronger	Gentleman 2 is stronger
acts	hold duel	1 wins, 2 injured (w_1)	2 wins, 1 injured (w_2)
	cancel duel	status quo (q)	

outcomes

	belief		utility function		
	1 stronger	2 stronger	w_1	w_2	q
Gentleman 1	90%	10%	1	-3	0
Gentleman 2	10%	90%	-4	1	0

Example due to Gilboa, Samet, and Schmeidler (2004)

Belief-averaging and relative utilitarianism ($BARU$): average beliefs and sum up $(0, 1)$-normalized utility functions

Theorem. (informal) $BARU$ is the only anonymous aggregation function satisfying **restricted monotonicity** and **independence of redundant acts**

		states	
		Gentleman 1 is stronger	Gentleman 2 is stronger
acts	hold duel	1 wins, 2 injured (w_1)	2 wins, 1 injured (w_2)
	cancel duel	status quo (q)	
		outcomes	

	belief		utility function		
	1 stronger	2 stronger	w_1	w_2	q
Gentleman 1	90%	10%	1 1	0 -3	$.75$ 0
Gentleman 2	10%	90%	0 -4	1 1	$.8$ 0
BARU	50%	50%	1	1	1.55

$$\mathbb{E}(\text{cancel duel}) = 1.55 > 1 = \mathbb{E}(\text{hold duel})$$

	EU Preferences	SEU Preferences
Single-profile	linear aggregation of utilities (Harsanyi, 1955)	linear aggregation of beliefs and utilities (Gilboa et al., 2004)
Multi-profile	relative utilitarianism (Dhillon and Mertens, 1999)	belief-averaging and relative utilitarianism

$(\Omega, \mathcal{E})$ measurable space of states

$(O, \mathcal{F})$ measurable space of outcomes

$\{f\colon \Omega \to O\colon f$ measurable$\}$ **acts**

$\mathcal{R}$ **SEU preference relations** over acts

- **belief**: non-atomic probability measure on $(\Omega, \mathcal{E})$
- **utility function**: measurable and bounded function $O \to \mathbb{R}$

$$f \succcurlyeq g \longleftrightarrow \int_{\Omega} (u \circ f) d\pi \geq \int_{\Omega} (u \circ g) d\pi$$

choose u with $\inf_{o \in O} u(o) = 0$ and $\sup_{o \in O} u(o) = 1$

This talk: all preference relations are SEU

Preference profile: tuple of SEU preferences $P = (\succcurlyeq_i)_{i \in I} \in \mathcal{R}^I$ for finite $I \subset \mathbb{N}$

Aggregation function: mapping every preference profile to a collective preference relation

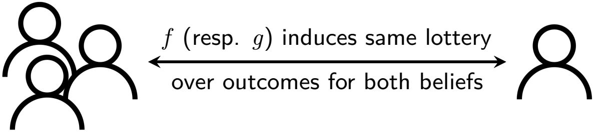
$$F\colon \bigcup_{I \subset \mathbb{N}} \mathcal{R}^I \to \mathcal{R}$$

Notation: $\pi_{\succcurlyeq}$ and $u_{\succcurlyeq}$ the belief and utility function representing $\succcurlyeq$ (π_i and u_i instead of $\pi_{\succcurlyeq_i}$ and $u_{\succcurlyeq_i}$)

Restricted Monotonicity

For all $P \in \mathcal{R}^I$, $j \notin I$

$f \sim_{F(P)} g$ $\qquad$ $f \succsim_j g$



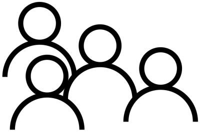
$$f_* \pi_{F(P)} = f_* \pi_j$$

$$g_* \pi_{F(P)} = g_* \pi_j$$

Restricted Monotonicity

For all $P \in \mathcal{R}^I$, $j \notin I$

$$f \succcurlyeq_{F(P_{\mathbf{+j}})} g$$



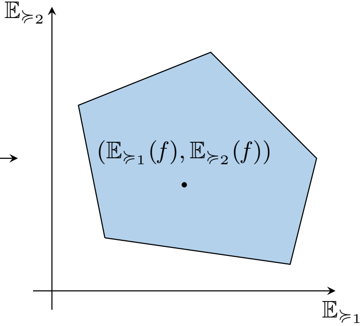
$$f_* \pi_{F(P)} = f_* \pi_j$$

$$g_* \pi_{F(P)} = g_* \pi_j$$

Independence of redundant acts

A set of acts $\mathcal{A}$ is **co-redundant** for a profile P if for every act, there is some act in $\mathcal{A}$ so that every agent is indifferent between the two

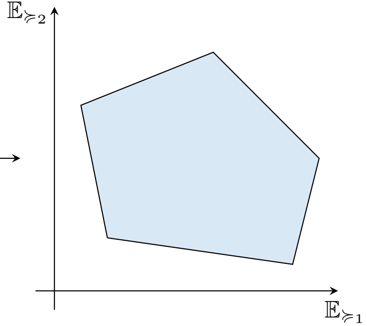
$P = (\succcurlyeq_1, \succcurlyeq_2)$: acts $\longrightarrow$



Independence of redundant acts

A set of acts $\mathcal{A}$ is **co-redundant** for a profile P if for every act, there is some act in $\mathcal{A}$ so that every agent is indifferent between the two

$P = (\succsim_1, \succsim_2)$: $\mathcal{A}$ $\longrightarrow$



If

$\mathcal{A}$ is **co-redundant** for P and P' and

every agent has the **same preferences over** $\mathcal{A}$ in P and P'

then $F(P)|_{\mathcal{A}} = F(P')|_{\mathcal{A}}$

Faithfulness

The collective preferences for any single-agent profile are those of the only agent

No belief imposition

No agent can dictate the collective belief

Continuity

Small changes in the agents' preferences lead to small changes in the collective preferences

Theorem. Every aggregation function F satisfying the above axioms is weighted belief-averaging and weighted utilitarian: there are $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^{\mathbb{N}}_{++}$ such that for all $P \in \mathcal{R}^I$, $F(P) = \succcurlyeq$ with

$$\pi_{\succcurlyeq} = \frac{1}{\sum_{i \in I} v_i} \sum_{i \in I} v_i \pi_i \qquad u_{\succcurlyeq} = \sum_{i \in I} w_i u_i$$

Corollary. The only aggregation function satisfying the above axioms and **anonymity** is belief-averaging and relative utilitarianism: for $P \in \mathcal{R}^I$, $BARU(P) = \succcurlyeq$ with

$$\pi_{\succcurlyeq} = \frac{1}{|I|} \sum_{i \in I} \pi_i \qquad u_{\succcurlyeq} = \sum_{i \in I} u_i$$

average of beliefs

sum of normalized utility functions

↓ restricted Pareto (Gilboa et al., 2004)

$$\pi_{F(P)} = \frac{1}{\sum_i v_i(P)} \sum_i v_i(P)\pi_i \qquad u_{F(P)} = \sum_i w_i(P) u_i$$

↓ restricted monotonicity

$$\sum_i v_i(\succsim_i)\pi_i \qquad \sum_i w_i(\succsim_i) u_i$$

↓ independence of redundant acts

$$\sum_i v_i \pi_i \qquad \sum_i w_i u_i$$

Assume $\pi_{F(P)} = \frac{1}{\sum_i v_i(P)} \sum_i v_i(P)\pi_i$ and $u_{F(P)} = \sum_i w_i(P) u_i$

Show: $w_i(P)$ is independent of P_{-i}

Idea:

Use **restricted monotonicity** to show that $\frac{w_i(P)}{w_j(P)}$ does not depend on $\succcurlyeq_k$, $k \neq i, j$

Consider $P = (\succcurlyeq_1, \succcurlyeq_2, \succcurlyeq_3)$ and $P_{-3} = (\succcurlyeq_1, \succcurlyeq_2)$

$$u_{F(P)} = \alpha u_{F(P_{-3})} + \beta u_3 = \alpha \left(\sum_{i=1}^{2} w_i(P_{-3}) u_i \right) + \beta u_3$$

$$= \sum_{i=1}^{3} w_i(P) u_i$$

The relative weights of agents 1 & 2 are equal in P and P_{-3}:
$w_1(P) = \alpha w_1(P_{-3})$ and $w_2(P) = \alpha w_2(P_{-3})$

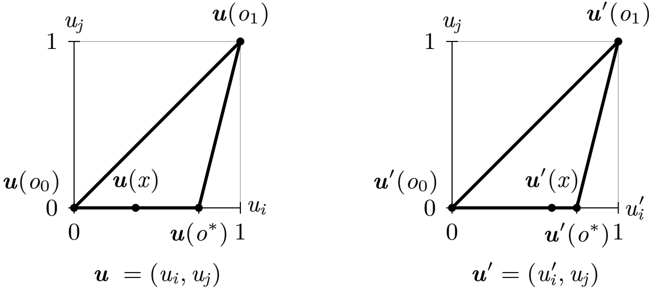
Can choose w so that $w_1(P) = w_1(P_{-3})$

Assume v_i is constant and w_i is independent of π_i
$\implies w_i \colon U \to \mathbb{R}_{++}$

Show: $w_i(u_i)$ is constant

Ideas:

- Use **IRA** to show that certain changes to u_i do not change $w_i(u_i)$
- Construct a path between any two utility functions along which w_i is constant



Belief-averaging and relative utilitarianism

$$\pi_{F(P)} = \frac{1}{|I|} \sum_{i \in I} \pi_i \qquad\qquad u_{F(P)} = \sum_{i \in I} u_i$$

Ex-ante relative utilitarianism (Sprumont, 2019)

$$f \succcurlyeq_{F(P)} g \quad \longleftrightarrow \quad \sum_{i \in I} \mathbb{E}_{\succcurlyeq_i}(f) \geq \sum_{i \in I} \mathbb{E}_{\succcurlyeq_i}(g)$$

Geometric-linear aggregation (Dietrich, 2019)

$$(\pi_{F(P)})(\omega) \sim \prod_{i \in I} (\pi_i(\omega))^{v_i(u_P)} \qquad\qquad u_{F(P)} = \sum_{i \in I} w_i(u_P) u_i$$

Restricted Pareto/monotonicity is susceptible **"complementary ignorance"** (Mongin and Pivato, 2019)

Neutrality on outcomes

Beliefs and utilities are **aggregated separately**