Social Choice

Christoph Schottmüller

Outline

Motivation

- people have different preferences
- how should societal decisions be taken?
  - navigate conflicts of preferences
  - respecting preferences
- Examples:
  - political decisions and elections
  - a group of friends wants to go for drinks: how to aggregate the differing preferences over bars
  - aggregating votes of several judges in sports (boxing, figure skating etc.)
  - (expert) committees
  - a family deciding where to spend the summer holiday
  - ...

Social choice theory

- make ethical premises explicit
- derive solutions consistent with these premises
- normative (!)

Example: Majority voting

- society ($N > 2$ people) has to choose one of 2 alternatives/candidates (x and y)
- assumption for simplicity: everyone has a strict preference over alternatives
- majority voting:
  - $x \succeq_S y$ if at least $N/2$ people prefer x over y
  - $y \succeq_S x$ if at least $N/2$ people prefer y over x
- what normative premises underlie this *social welfare function*?

Some criteria (for 2 alternatives)

Anonymity

A social choice function is *anonymous* if the names of the agents do not matter, i.e. if a permutation of preferences across agents does not change the social preference.

Neutrality

A social welfare function is *neutral* if the names of the alternatives do not matter, i.e. the social preferences are reversed if we reverse the preferences of all agents.

Positive responsiveness

A social welfare function is *positively responsive* if the following holds: if one alternative, say x, is weakly socially preferred although $y \succ_i x$ for some $i \in \{1, \ldots, N\}$, then x is strictly socially preferred if we change i's preferences (without changing anyone else's preferences).

A first theorem

- didn't I claim that social choice starts with premises and then derives solutions?

May's Theorem

If there are two alternatives, a social welfare function satisfies anonymity, neutrality and positive responsiveness if and only if it is majority voting.

Proof sketch ("only if" for even N)

- Anonymity: only number of people preferring alternative x over y matters for $\succeq_S$.
- Neutrality: if $N/2$ people prefer x over y, then $x \approx_S y$.
- Positive responsiveness: if more than $N/2$ people prefer x over y, $x \succ_S y$ and vice versa.

Majority voting with more than 2 alternatives

- How to generalize majority voting with more than 2 alternatives?

Definition

An alternative x is a *Condorcet winner* if for any other alternative y a majority prefers x over y.

Example

A group of students want to tell the teacher their preferences over exam forms (open book, closed book, online exam). How to aggregate the preferences?

	best	middle	worst
Student 1	ob	oe	cb
Student 2	oe	cb	ob
Student 3	cb	ob	oe

Which alternative is Condorcet winner?

Model

- finite set $X = \{x_1, x_2, \ldots, x_K\}$ of alternatives
- $N \geq 2$ agents, each has a complete and transitive preference relation over X

Social preference relation

A social preference relation is a complete and transitive preference relation on the set X.

Social welfare function

A social welfare function assigns to each profile of preferences $(\succeq_1, \succeq_2, \ldots, \succeq_N)$ a social preference relation $\succeq_S$.

Examples: social welfare function

Are the following social welfare functions desirable?

- The preferences of agent 1 are the social preferences: $\succeq_S (\succeq_1, \succeq_2, \ldots, \succeq_N) = \succeq_1$
- Fixed social preference relation: $\succeq_S (\succeq_1, \succeq_2, \ldots, \succeq_N) = x_1 \succ_S x_2 \succ_S x_3 \succ_S \cdots \succ_S x_K$
- Borda Count:
  - turn every agent's preference order into points: the k most preferred alternative receives k points
  - for every alternative, sum the points it gets from all agents
  - order alternatives according to points

Borda and Olympic Ice Skating competition I

- judging in sports is similar to our problem
  - aggregation of several judges' rankings
- final 2002 Olympic figure skating competition
  - Slutskaya is the last skater to perform
  - at that moment: 1. Kwan, 2. Hughes, 3. ...
  - Slutskaya is doing well but not super and ends up second
  - who came first? who came third?

Borda and Olympic Ice Skating competition II

- say, first rank gives 3 points, second 2 and third 1

	Kwan	Hughes	Slutskaya
judge 1	2	3	1
judge 2	2	3	1
judge 3	1	2	3
judge 4	1	2	3
judge 5	3	1	2
judge 6	3	1	2
judge 7	3	1	2
Points			

Minimal (?) normative criteria

Weak Pareto principle (unanimity)

If $x \succ_i y$ for all $i = 1, 2, \ldots, N$, then $x \succ_S y$.

Non-dictatorship

There is no individual i such that $x \succeq_S y$ if and only if $x \succeq_i y$. (no matter what other agents preferences are)

Independence of irrelevant alternatives

Take two profiles of preferences $(\succeq_1, \succeq_2, \ldots, \succeq_N)$ and $(\succeq'_1, \succeq'_2, \ldots, \succeq'_N)$. If for every agent i the ranking of x and y is the same under $\succeq_i$ and $\succeq'_i$, then the social ranking of x and y must be the same under these two preference profiles.[a]

[a] More formally, let the two preference profiles be such that for all agents i $x \succeq_i y$ if and only if $x \succeq'_i y$. Then $x \succeq_S y$ if and only if $x \succeq'_S y$.

Arrow's impossibility theorem

Theorem

Let there be at least be 3 alternatives in X. There exists no social welfare function that satisfies all 3 criteria (weak Pareto principle, non-dictatorship and independence of irrelevant alternatives).

Proof is somewhat lengthy (see textbook)

Consequences of Arrow's theorem

- no social welfare function satisfies even minimal criteria
- we have to give up even some of these minimal criteria if we want to proceed!
- some ways to proceed:
  - pick only one alternative: no complete social ordering necessary
    - leads to similar result
  - domain restriction
    - we implicitly assumed that all preference profiles were possible (in the definition "social welfare function")
    - more positive results if we can rule out certain preferences
  - cardinal utility
    - we only looked at orderings not at intensity of preference
    - assuming that there is something like intensity of preferences *and this intensity is comparable across agents* helps to aggregate preferences but is a questionable assumption

Domain restriction: Single peaked preferences I

- imagine alternatives are ordered on a real line $x_1 < x_2 < \cdots < x_K$
- assumptions:
  - common ordering of alternatives
  - everyone has a most preferred alternative
  - of two "too high" (or "too low") alternatives, an agent prefers the one closer to his most preferred alternative
  - for simplicity: odd number N of agents
- more precisely:
  - each agent i has a most preferred alternative $x^*(i) \in \{x_1, x_2, \ldots, x_K\}$
  - if $x_k, x_m > x^*(i)$, then $x_k \succ_i x_m$ if and only if $x_k < x_m$
  - if $x_k, x_m < x^*(i)$, then $x_k \succ_i x_m$ if and only if $x_k > x_m$
- if we represent preferences by utility function, this function is "single peaked"

Domain restriction: Single peaked preferences II

Median agent for single peaked preferences

An agent i is a *median agent* if
(i) there are at least $N/2$ agents with most preferred alternatives weakly above $x^*(i)$ and
(ii) there are at least $N/2$ agents with most preferred alternatives weakly below $x^*(i)$.

Note: a median agent always exists.

Domain restriction: Single peaked preferences II

Proposition

Let preferences be single peaked and i be a median agent, then $x^*(i)$ is a Condorcet winner.

Proof

- Consider a pairwise majority vote between $x^*(i)$ and $x_m > x^*(i)$.
- Consider a pairwise majority vote between $x^*(i)$ and $x_m < x^*(i)$.

Domain restriction: Single peaked preferences III

- consider pairwise majority voting between arbitrary alternatives, i.e. say x_k is socially preferred to x_m if x_k wins in a majority vote over x_k and x_m

Proposition

If preferences are single peaked, pairwise majority voting induces a social welfare function.

Proof

to show: resulting preferences are complete and transitive

Cardinal utility I

Reminder:

Representation by a utility function

A complete preference relation $\succeq$ over a set X is represented by the utility function $u : X \to \Re$ if and only if

$$x \succeq y \quad \Leftrightarrow \quad u(x) \geq u(y).$$

If u represents $\succeq$, then $\psi(u)$ also represents $\succeq$ where $\psi : \Re \to \Re$ is an arbitrary strictly increasing function.

Cardinal utility II

- suppose we have 2 agents and $x \succ_1 y$ while $y \succ_2 x$
- we choose utility functions for the two agents
  - $u_1(x) = 3,\ u_1(y) = 1$
  - $u_2(x) = 0,\ u_2(y) = 1$
- which alternative should society prefer?

Cardinal utility II

- if we assign meaning to utility, social welfare function is not invariant to strictly monotone transformations
- allows to get around Arrow's impossibility theorem
- problem: choice of specific agent utility functions implicitly makes normative judgments beyond our criteria
- for now:
    - accept some given utility functions u
    - let welfare depend on the utilities of the agents and be represented by a function $W : \Re^N \to \Re$ that aggregates agent utilities into "welfare"
        - we abuse notation and call W also "social welfare function"
    - what are reasonable choices for W? what normative judgments are expressed by the choice of W?

Cardinal utility III

Pareto dominance

Alternative x is *Pareto dominated* by alternative y if and only if $y \succeq_i x$ for all agents $i = 1, .., N$ and $y \succ_i x$ for at least one agent.

Pareto efficiency

An alternative x is *Pareto efficient* if there is no alternative y that Pareto dominates x.

Cardinal utility IV

Proposition

If social welfare function W is strictly increasing, then Pareto dominating alternatives are socially preferred to the alternatives they dominate.

Proof

- let W be strictly increasing and x Pareto dominate y

Cardinal utility V: Rawlsian welfare

$$W_{Rawls}(u_1, \ldots, u_N) = \min[u_1, \ldots, u_N]$$

- W_{Rawls} is strictly increasing $\Rightarrow$ satisfies Pareto criterion
- W_{Rawls} is anonymous
- W_{Rawls} is "utility level invariant":
  - social preferences remain the same if we transform all agent's utility using *the same* strictly increasing transformation
- W_{Rawls} satisfies "Hammond Equity":
  - take two utility vectors $(\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_N)$ and $(\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_N)$ and suppose $\bar{u}_i = \hat{u}_i$ for all i except j and k
  - suppose further $\bar{u}_j < \hat{u}_j < \hat{u}_k < \bar{u}_k$
  - Hammond equity states that then $W(\hat{u}) > W(\bar{u})$

Cardinal utility VI: Rawlsian welfare

Proposition

A strictly increasing and continuous social welfare function W satisfies Hammond equality if and only if it can take the Rawlsian form $W_{Rawls}(u_1, \ldots, u_N) = \min[u_1, \ldots, u_N]$.

- $\approx$ Rawlsian welfare is equivalent to Pareto criterion $+$ Hammond equity

Proof

see Jehle and Reny (2011), section 6.3.1

Cardinal utility VII: Utilitarian welfare

$$W_{ut}(u_1, \ldots, u_N) = \sum_{i=1}^{N} u_i$$

- most common form of welfare function (sometimes with individual weights)
- W_{ut} is strictly increasing $\Rightarrow$ satisfies Pareto criterion
- W_{ut} is anonymous (not true if weights are used)
- W_{ut} is "utility-difference invariant"
  - social preferences are the same if we transform all agents utility using the transformation $\psi_i(u_i) = a_i + bu_i$

Cardinal utility VIII: Utilitarian welfare

Proposition

A strictly increasing and continuous social welfare function W satisfies anonymity and utility-difference invariance if and only if it can take the utilitarian form $W_{ut} = \sum_{i=1}^{N} u_i$.

Proof

see Jehle and Reny (2011), section 6.3.2

Cardinal utility IX: the veil of ignorance I

- thought experiment
  - you will be one of the agents in society
  - you have to decide which alternative to choose
  - you do not know which agent you are going to be
  - some people have argued that whatever a "fair-minded" person would choose in this hypothetical situation is a good societal decision

Cardinal utility X: the veil of ignorance II

- Harsanyi:
  - my chance of being agent i is $1/N$
  - my choice should maximizes the expected utility $\sum_{i=1}^{N}(1/N)u_i(x)$
  - $\rightarrow$ utilitarian welfare
- Rawls:
  - I do not know who I am going to be and there is no basis for assigning probabilities.
  - risk aversion implies maximizing the worst case utility
  - $\rightarrow$ Rawlsian welfare
- Arrow:
  - Rawls makes a mistake as he assumes not risk aversion but *infinite* risk aversion, i.e. risk aversion does *not* imply maximizing worst case utility.

Manipulability I

- so far: preferences of all players are known
- problem: aggregation
- what if everyone knows his preferences privately?
  - ask for preferences
  - aggregate
- additional problem: gaming the system by misreporting preferences!
- result due to Gibbard and Satterthwaite:
  If there are at least three alternatives and a social welfare function is (i) Pareto efficient and (ii) creates no gaming possibilities, then it is dictatorial.

Manipulability II

- one example for manipulability

Example: Borda count

	most preferred	middle preferred	least preferred
Agent 1	x	y	z
Agent 2	y	x	z
Agent 3	y	x	z
Points			

Could agent 1 manipulate the social preference relation by misrepresenting his own preferences? Would he want to do so?

- to discuss such topics properly:
  extend decision and game theory to incomplete information
  - that's what we will do in the coming weeks!