Lecture 4: Arrow's Theorem and Related Results

Eric Pacuit

ILLC, University of Amsterdam

staff.science.uva.nl/~epacuit

epacuit@science.uva.nl

Lecture Date: March 2 , 2006

Caput Logic, Language and Information: Social Software

staff.science.uva.nl/~epacuit/caputLLI.html

First, a proof of Kuhn's Theorem

Subgame Perfect Equilibrium

The **subgame of the extensive game with perfect information** $G = \langle N, H, P, (\succeq_i) \rangle$ that follows history h is the game

$$G_h = \langle N, H_h, P_n, (\succeq_i^h) \rangle$$

where

- H_h is the set of sequences of h' of actions for which $(h, h') \in H$
- P_h is defined $P|_h(h') = P(h, h')$ for each $h' \in H_h$
- $\succeq_i^h$ is define $h' \succeq_i^h h''$ iff $(h, h') \succeq_i (h, h'')$
- Given a strategy σ, let σ^h denote the history induced on G_h by σ
- Let O_h denote the outcome function of G_h.

Subgame Perfect Equilibrium

A strategy profile σ is a subgame perfect equilibrium of a game G if for every $i \in N$ and every nonterminal history h for which $P(h) = i$ we have

$$O_h(\sigma^h_{-i}, \sigma^h_i) \succeq^h_i O_h(\sigma^h_{-i}, \tau_i)$$

for every strategy τ_i of player i in the subgame G_h.

That is, σ^* is a subgame perfect equilibrium if σ^h is a Nash Equilibrium in every subgame of G.

Subgame Perfect Equilibrium

Kuhn's Theorem Every finite extensive game with perfect informaiton has a subgame perfect equilibrium.

Subgame Perfect Equilibrium

An extensive game is a **finite horizon** game if the length of each history is finite.

Lemma (The one deviation property) Let G be a finite horizon extensive game. The strategy profile σ is a subgame perfect equilibrium of G iff for every player $i \in N$ and every history $h \in H$ for which $P(h) = i$ we have

$$O_h(\sigma_{-i}^h, \sigma_i^h) \succeq_i^h O_h(\sigma_{-i}^h, \tau_i)$$

for every strategy τ_i of player i in the subgame G_h that differs from σ_i^h only in the action it prescribes after the initial history of G_h.

Quick Review

Quick Review

- **Lecture 1:** What is social software?

  Goal of the course: study mathematical tools relevant for the analysis and design of *social procedures*

Quick Review

- **Lecture 1:** What is social software?

  Goal of the course: study mathematical tools relevant for the analysis and design of *social procedures*

- **Lecture 2:** Expected Utility Theorem

  Representing **individual** decision makers

Quick Review

- **Lecture 1:** What is social software?

  Goal of the course: study mathematical tools relevant for the analysis and design of *social procedures*

- **Lecture 2:** Expected Utility Theorem

  Representing **individual** decision makers

- **Lecture 3:** Describing strategic situations (extensive and strategic games)

  Individual decision makers *interacting*

Today's Lecture: **What is the logic of group choice?**

Our goal this semester is to develop and study frameworks for studying *social* procedures

Today's Lecture: What is the logic of group choice?

Our goal this semester is to develop and study frameworks for studying *social* procedures

I.e., (game-theoretic) mechanisms intended to solve some societal problem (select a candidate to govern, divide a cake, divide a set of goods among a group of people, etc.)

Today's Lecture: **What is the logic of group choice?**

Our goal this semester is to develop and study frameworks for studying *social* procedures

I.e., (game-theoretic) mechanisms intended to solve some societal problem (select a candidate to govern, divide a cake, divide a set of goods among a group of people, etc.)

Many social procedure can be described (very generally) as follows:

Today's Lecture: **What is the logic of group choice?**

Our goal this semester is to develop and study frameworks for studying *social* procedures

I.e., (game-theoretic) mechanisms intended to solve some societal problem (select a candidate to govern, divide a cake, divide a set of goods among a group of people, etc.)

Many social procedure can be described (very generally) as follows:

1. Allow each agent to represent its personal preference over a set of possible outcomes

Today's Lecture: What is the logic of group choice?

Our goal this semester is to develop and study frameworks for studying *social* procedures

I.e., (game-theoretic) mechanisms intended to solve some societal problem (select a candidate to govern, divide a cake, divide a set of goods among a group of people, etc.)

Many social procedure can be described (very generally) as follows:

1. Allow each agent to represent its personal preference over a set of possible outcomes
2. Somehow combine this information in order to find a social preference (or select a particular outcome, called the social choice)

Today's Lecture: **What is the logic of group choice?**

As we saw with the Condorcet Paradox, group preferences (generated by majority rule) do not necessarily have the same properties as the individual preferences.

Today's Lecture: What is the logic of group choice?

As we saw with the Condorcet Paradox, group preferences (generated by majority rule) do not necessarily have the same properties as the individual preferences.

This raises a question: *Is it always possible to find a procedure that will satisfy certain natural properties?*

Today's Lecture: What is the logic of group choice?

As we saw with the Condorcet Paradox, group preferences (generated by majority rule) do not necessarily have the same properties as the individual preferences.

This raises a question: *Is it always possible to find a procedure that will satisfy certain natural properties?*

- A Possibility Result: May's Theorem
- An Impossibility Result: Arrow's Theorem

Reminder: Agent Preference

Let X be a set of alternatives (usually finite) and $\mathcal{A} = \{1, \ldots, n\}$ a set of agents. For $x, y \in X$ and $i \in \mathcal{A}$

Reminder: Agent Preference

Let X be a set of alternatives (usually finite) and $\mathcal{A} = \{1, \ldots, n\}$ a set of agents. For $x, y \in X$ and $i \in \mathcal{A}$

xP_iy: i **prefers** x over y

Reminder: Agent Preference

Let X be a set of alternatives (usually finite) and $\mathcal{A} = \{1, \ldots, n\}$ a set of agents. For $x, y \in X$ and $i \in \mathcal{A}$

xP_iy: i **prefers** x over y

xI_iy: i is **indifferent** between x and y

$$xI_iy \quad \text{iff} \quad \text{not } (xP_iy) \quad \text{and} \quad \text{not } (yP_ix)$$

Reminder: Agent Preference

Let X be a set of alternatives (usually finite) and $\mathcal{A} = \{1, \ldots, n\}$ a set of agents. For $x, y \in X$ and $i \in \mathcal{A}$

xP_iy: i **prefers** x over y

xI_iy: i is **indifferent** between x and y

$$xI_iy \text{ iff } \text{ not } (xP_iy) \quad \text{and} \quad \text{not } (yP_ix)$$

xR_iy: i **weakly prefers** x over y

$$xR_iy \text{ iff } xP_iy \quad \text{or} \quad yI_ix$$

Reminder: Agent Preference

Given a set of agent $\mathcal{A} = \{1, \ldots, n\}$.

Let $\mathcal{P}$ be the set of all possible preferences (over a set of outcomes X).

A **preference profile** is a vector of preferences $\vec{P} = \langle P_1, \ldots, P_n \rangle$.

Let $\vec{\mathcal{P}} = \Pi_{i \in \mathcal{A}} \mathcal{P}$ be the set of all possible preference profiles.

A **social welfare function** (SWF) is a function $\mathsf{F} : \vec{\mathcal{P}} \rightarrow X$.

A **social welfare relation** (SWR) is a relation $\mathsf{R} : \vec{\mathcal{P}} \rightarrow \mathcal{P}$

Reminder: Agent Preference

Given a set of agent $\mathcal{A} = \{1, \ldots, n\}$.

Let $\mathcal{P}$ be the set of all possible preferences (over a set of outcomes X).

A **preference profile** is a vector of preferences $\vec{P} = \langle P_1, \ldots, P_n \rangle$.

Let $\vec{\mathcal{P}} = \Pi_{i \in \mathcal{A}} \mathcal{P}$ be the set of all possible preference profiles.

A **social welfare function** (SWF) is a function $\mathsf{F} : \vec{\mathcal{P}} \rightarrow X$.

A **social welfare relation** (SWR) is a relation $\mathsf{R} : \vec{\mathcal{P}} \rightarrow \mathcal{P}$

Simple Majority Rule

- Suppose that there are n individuals and **two** alternatives x and y

Simple Majority Rule

- Suppose that there are n individuals and **two** alternatives x and y

- For each i there is a variable $D_i \in \{-1, 0, 1\}$ where

$$D_i = \begin{cases} -1 & \text{if } \; yP_ix \\ 0 & \text{if } \; xI_iy \\ 1 & \text{if } \; xP_iy \end{cases}$$

Simple Majority Rule

- Suppose that there are n individuals and **two** alternatives x and y

- For each i there is a variable $D_i \in \{-1, 0, 1\}$ where

$$D_i = \begin{cases} -1 & \text{if } \; yP_ix \\ 0 & \text{if } \; xI_iy \\ 1 & \text{if } \; xP_iy \end{cases}$$

- $\mathsf{F} : \{-1, 0, 1\}^n \rightarrow \{-1, 0, 1\}$ is a **group decision function**

Example: Simple Majority Rule

For $k \in \{-1, 0, 1\}$, let

$$N_k(D_1, \ldots, D_n) = |\{i \mid D_i = k\}|$$

Let $\vec{D} = \langle D_1, \ldots, D_n \rangle$

F_{maj} is a **simple majority decision method** iff

$$\mathsf{F}_{maj}(\vec{D}) = \begin{cases} -1 & \text{if } N_1(\vec{D}) - N_{-1}(\vec{D}) < 0 \\ 0 & \text{if } N_1(\vec{D}) - N_{-1}(\vec{D}) = 0 \\ 1 & \text{if } N_1(\vec{D}) - N_{-1}(\vec{D}) > 0 \end{cases}$$

Properties of group decision funcitons

A group decision function F is

Properties of group decision funcitons

A group decision function F is

- **Decisive** if F is a total function

Properties of group decision funcitons

A group decision function F is

- **Decisive** if F is a total function

- **Symmetric** if $\mathsf{F}(D_1, \ldots, D_n) = \mathsf{F}(D_{\pi(1)}, \ldots, D_{\pi(n)})$ for all permutations π. I.e., F is symmetric in all of its arguments.

Properties of group decision funcitons

A group decision function F is

- **Decisive** if F is a total function

- **Symmetric** if $\mathsf{F}(D_1, \ldots, D_n) = \mathsf{F}(D_{\pi(1)}, \ldots, D_{\pi(n)})$ for all permutations π. I.e., F is symmetric in all of its arguments.

- **Neutral** if $\mathsf{F}(-D_1, \ldots, -D_n) = -\mathsf{F}(D_1, \ldots, D_n)$

Properties of group decision funcitons

A group decision function F is

- **Decisive** if F is a total function

- **Symmetric** if $\mathsf{F}(D_1, \ldots, D_n) = \mathsf{F}(D_{\pi(1)}, \ldots, D_{\pi(n)})$ for all permutations π. I.e., F is symmetric in all of its arguments.

- **Neutral** if $\mathsf{F}(-D_1, \ldots, -D_n) = -\mathsf{F}(D_1, \ldots, D_n)$

- **Positively Responsive** if $\mathsf{F}(D_1, \ldots, D_n) = 0$ or 1, and $D'_i = D_i$ for all $i \neq i_0$, and $D'_{i_0} > D_{i_0}$, then $\mathsf{F}(D'_1, \ldots, D'_n) = 1$

May's Theorem (1952)

Theorem A group decision function is the method of simple majority decision if and only if it is always decisive, symmetric, neutral and positively responsive

Digression: May's Theorem for an infinite population

Assume $\mathcal{A}$ is a **countable** set. Can we prove May's Theorem in this setting?

What is the majority of an infinite set?

All properties of F generalize to the infinite setting except symmetry (Why?).

See (Mark Fey, 2001) and (P and Salame, 2005) for a discussion of majorities of an infinite set.

The General Situation

In May's Theorem, the agents are making a single binary choice between two alternatives.

What about more general situations?

- Agents choose between between more than two alternatives.
- There are multiple interconnected propositions on which simultaneous decisions are to be made.

The General Situation

In May's Theorem, the agents are making a single binary choice between two alternatives.

What about more general situations?

- Agents choose between between more than two alternatives.
- There are multiple interconnected propositions on which simultaneous decisions are to be made. (Later this semester)

Arrow's Theorem

Social Choice & Individual Values by Kenneth J. Arrow (1951)

Also,

"Three Brief Proofs of Arrow's Impossibility Theorem." John Geanakoplos. *Economic Theory*, **26**, 2005

Social Choice and The Mathematics of Manipulation by Alan D. Taylor (2005)

Mathematical Methods in Economics and Social Choice by Norman Shofield (2002)

Arrow's Theorem

Let X be a finite set of preferences with *at least three elements.*

Assume each agent has a transitive and complete preference over X (ties are allowed).

Arrow's Theorem

Let X be a finite set of preferences with *at least three elements.*

Assume each agent has a transitive and complete preference over X (ties are allowed).

> In the present study the objects of choice are social states. The most precise definition of a social state would be a complete description of the amount of each type of commodity in the hands of each individual , the amount of labor to be supplied by each individual , the amount of each productive resource invested in each type of productive activity . . . and the erection of statues to famous men. (Arrow, 1951)

Arrow's Theorem: Unanimity

If each agent ranks x above y, then so does the social welfare function

If for each $i \in \mathcal{A}$, xP_iy then $x\mathsf{R}(\vec{P})y$

Arrow's Theorem: Independence of Irrelevant Alternatives

The social relative ranking (higher, lower, or indifferent) of two alternatives x and y depends only the relative rankings of x and y for each individual.

$x\mathsf{R}(\vec{P})y$ iff $x\mathsf{R}(\vec{P}')y$ whenever for each $i \in \mathcal{A}$, xP_iy iff xP'_iy.

Arrow's Theorem: Dictatorship

There is an individual $d \in \mathcal{A}$ such that the society strictly prefers x over y whenever d strictly prefers x over y.

There is a $d \in \mathcal{A}$ such that $x \mathsf{R}(\vec{P}) y$ whenever $x P_d y$.

Arrow's Theorem

Theorem (Arrow, 1951) Any social welfare function that respects transitivity, independence of irrelevant alternatives and unanimity is a dictatorship.

Arrow's Theorem: Proof 1 (Geanakoplos, 2005)

Arrow's Theorem: Proof 1 (Geanakoplos, 2005)

1. For any profile in which every agent puts alternative x at the very top or the very bottom of his ranking, then society must as well.

Arrow's Theorem: Proof 1 (Geanakoplos, 2005)

1. For any profile in which every agent puts alternative x at the very top or the very bottom of his ranking, then society must as well.

2. There exists a voter d that by changing his vote at some profile he can move x from the bottom of the social ranking to the top.

Arrow's Theorem: Proof 1 (Geanakoplos, 2005)

1. For any profile in which every agent puts alternative x at the very top or the very bottom of his ranking, then society must as well.

2. There exists a voter d that by changing his vote at some profile he can move x from the bottom of the social ranking to the top.

3. This agent d is a dictator over any pair y, z not involving x

Arrow's Theorem: Proof 1 (Geanakoplos, 2005)

1. For any profile in which every agent puts alternative x at the very top or the very bottom of his ranking, then society must as well.

2. There exists a voter d that by changing his vote at some profile he can move x from the bottom of the social ranking to the top.

3. This agent d is a dictator over any pair y, z not involving x

4. This agent d is also a dictator over every pair x, z.

Arrow's Theorem: Proof 1 (Geanakoplos, 2005)

1. For any profile in which every agent puts alternative x at the very top or the very bottom of his ranking, then society must as well.

2. There exists a voter d that by changing his vote at some profile he can move x from the bottom of the social ranking to the top.

3. This agent d is a dictator over any pair y, z not involving x

4. This agent d is also a dictator over every pair x, z.

Arrow's Theorem: Proof 2

For a set X, let

$S(X)$ is the set of **strict preference relations** (irreflexive and asymmetric)

$T(X)$ is the set of **strict partial orders** (transitive strict preference)

$O(X)$ is the set of **weak orders** (negatively transitive strict preference)

P is negatively transitive iff $\forall x, y, z((\neg xPy \wedge \neg yPz) \rightarrow \neg yPx)$

Arrow's Theorem: Decisive Coalition

A set $M \subseteq \mathcal{A}$ is **decisive** under R iff for any profile $\vec{P}$ such that xP_iy for all $i \in M$ then $x\mathsf{R}(\vec{P})y$.

Let $\mathcal{D}_{\mathsf{R}}$ be the set of all decisive coalitions under the social welfare rule R.

Arrow's Theorem: Decisive Coalition

A set $M \subseteq \mathcal{A}$ is **decisive** under R iff for any profile $\vec{P}$ such that xP_iy for all $i \in M$ then $x\mathsf{R}(\vec{P})y$.

Let $\mathcal{D}_{\mathsf{R}}$ be the set of all decisive coalitions under the social welfare rule R.

Lemma If $\mathsf{R} : T(X)^n \to T(X)$ and $\mathcal{A}, M_1, M_2$ all belong to $\mathcal{D}_{\mathsf{R}}$, then $M_1 \cap M_2 \in \mathcal{D}_{\mathsf{R}}$

Arrow's Theorem: Proof 2

Lemma If $\mathsf{R} : T(X)^n \to T(X)$ and R satisfies the unanimity then $\mathcal{D}_{\mathsf{R}}$ is a filter:

1. If $M_1 \subseteq M_2$ and $M_1 \in \mathcal{D}_{\mathsf{R}}$ then $M_2 \in \mathcal{D}_{\mathsf{R}}$
2. $\mathcal{A} \in \mathcal{D}_{\mathsf{R}}$ and $\emptyset \notin \mathcal{D}_{\mathsf{R}}$
3. If $M_1, M_2 \in \mathcal{D}_{\mathsf{R}}$ then $M_1 \cap M_2 \in \mathcal{D}_{\mathsf{R}}$.

Arrow's Theorem: Proof 2

Lemma If $\mathsf{R} : T(X)^n \to T(X)$ and R satisfies the unanimity then $\mathcal{D}_\mathsf{R}$ is a filter:

1. If $M_1 \subseteq M_2$ and $M_1 \in \mathcal{D}_\mathsf{R}$ then $M_2 \in \mathcal{D}_\mathsf{R}$
2. $\mathcal{A} \in \mathcal{D}_\mathsf{R}$ and $\emptyset \notin \mathcal{D}_\mathsf{R}$
3. If $M_1, M_2 \in \mathcal{D}_\mathsf{R}$ then $M_1 \cap M_2 \in \mathcal{D}_\mathsf{R}$.

Lemma If $\mathsf{R} : T(X)^n \to O(X)$ and $\mathcal{A} \in \mathcal{D}_\mathsf{R}$ and $M \subseteq \mathcal{A}$ and $M \notin \mathcal{D}_\mathsf{R}$ then $\mathcal{A} - M \in \mathcal{D}_\mathsf{R}$

Arrow's Theorem: Proof 2

Lemma If $\mathsf{R} : T(X)^n \to T(X)$ and R satisfies the unanimity then $\mathcal{D}_{\mathsf{R}}$ is a filter:

1. If $M_1 \subseteq M_2$ and $M_1 \in \mathcal{D}_{\mathsf{R}}$ then $M_2 \in \mathcal{D}_{\mathsf{R}}$
2. $\mathcal{A} \in \mathcal{D}_{\mathsf{R}}$ and $\emptyset \notin \mathcal{D}_{\mathsf{R}}$
3. If $M_1, M_2 \in \mathcal{D}_{\mathsf{R}}$ then $M_1 \cap M_2 \in \mathcal{D}_{\mathsf{R}}$.

Lemma If $\mathsf{R} : T(X)^n \to O(X)$ and $\mathcal{A} \in \mathcal{D}_{\mathsf{R}}$ and $M \subseteq \mathcal{A}$ and $M \notin \mathcal{D}_{\mathsf{R}}$ then $\mathcal{A} - M \in \mathcal{D}_{\mathsf{R}}$

Arrow's Impossibility Theorem If $\mathsf{R} : O(X)^n \to O(X)$ and $\mathcal{A} \in \mathcal{D}_{\mathsf{R}}$, with $|\mathcal{A}|$ finite, then there exists a $i \in \mathcal{A}$ such that $\{i\} \in \mathcal{D}_{\mathsf{R}}$.

Next week: Examples of social procedures (fair division algorithms)

Next^2 week: Guest Lecture: K. Apt

Next^3 week: From reasoning about programs to reasoning about social procedures