Notes on the Proof of Arrow's Theorem

Eric Pacuit*

March 24, 2014

The candidates and voters

- X is a set of candidates and N is a set of voters.
- $|X| > 2$ (there are more than 2 candidates) and $|N| = n$ (there are finitely many voters)

Preferences

- A **preference relation** on X is a relation $R \subseteq X \times X$ such that for all $a, b \in X$:
  - **Reflexivity**: $a\ R\ a$
  - **Transitivity**: if $a\ R\ b$ and $b\ R\ c$, then $a\ R\ c$
  - **Connectedness**: $a\ R\ b$ or $b\ R\ a$
- Let R be a preference relation, define two preference relations:
  - **Strict Preference**: $a\ P\ b\ := \ a\ R\ b$ and $b \not R\ a$
  - **Indifference Relation**: $a\ I\ b\ := \ a\ R\ b$ and $b\ R\ a$
- Preference, Strict Preference and Indifference relations satisfy the following properties:
  - Strict preference P is a strict order (it is transitive and irreflexive: for all $a \in X$, $a \not P\ a$)
  - Indifference I is an equivalence relation (reflexive, transitive and symmetric: for all $a, b \in X$, if $a\ I\ b$, then $b\ I\ a$)
  - Trichotomy: for all $a, b \in X$, either $a\ P\ b$ or $b\ P\ a$ or $a\ I\ b$
  - Absorption: for all $a, b, c, d \in X$, if ($a\ I\ b$ and $b\ P\ c$ and $c\ I\ d$), then $a\ P\ d$

*epacuit@umd.edu

Profiles

- A profile p for the voters N, is a sequence of preference orderings of length n, I.e., a profile is an element of $O(X)^n$, we denote profile $p \in O(X)^n$ as follows $p = (R_1, \ldots, R_n)$.
- Given a profile p, let p_i denote i's preference ordering in p and $p_i^>$ denote i's strict preference ordering.
- For $Y \subseteq X$, let $p|_Y$ be the profile of preference orderings *restricted to the candidates in* Y.
- Let $U \subseteq N$ be a set of voters and $a, b \in X$ and $p \in O(X)^n$ a profile, let

  $$a\ p_U^>\ b \text{ iff for all } i \in U,\ a\ p_i^>\ b$$

  I.e., $a\ p_U^>\ b$ means all voters in U strictly prefer a over b.
- Let $\mathcal{D}$ be the set of *possible* profiles (i.e., $\mathcal{D} \subseteq O(X)^n$)
- Suppose that $U \subseteq N$ is a set of voters. For $a, b \in X$, let

  $$U_{ab} = \{p \in \mathcal{D} \mid a\ p_U^>\ b \text{ and } b\ p_{U^C}^>\ a\}$$

  where $U^C = \{i \in N \ \mid i \notin U\}$ is the complement of U.

Example: Suppose that there are three candidates $X = \{a, b, c\}$ and three voters $\{1, 2, 3\}$. Let p, q, r be the following three profiles (each voter has a strict preference over the candidates with the most preferred candidates at the top of the list):

$$p: \begin{array}{ccc} 1 & 2 & 3 \\ \hline a & b & c \\ b & c & a \\ c & a & b \end{array} \qquad q: \begin{array}{ccc} 1 & 2 & 3 \\ \hline a & b & a \\ b & a & b \\ c & c & c \end{array} \qquad r: \begin{array}{ccc} 1 & 2 & 3 \\ \hline b & b & a \\ a & c & b \\ c & a & c \end{array}$$

Finally, let s be the profile $\begin{array}{ll} 1: & a\ I\ b\ P\ c \\ 2: & a\ P\ b\ I\ c \\ 3: & b\ P\ c\ I\ a \end{array}$

You should verify that the following are true:

- For $U = \{1, 3\}$: $a\ p_U^>\ b$, $\quad a\ q_U^>\ b$, $\quad a\ q_U^>\ c$, $\quad b\ q_U^>\ c$, $\quad a\ r_U^>\ c$, $\quad b\ r_U^>\ c$, and $\quad b\ s_U^>\ c$
- $a\ q_N^>\ c$, $\quad b\ q_N^>\ c$, $\quad$ and $\quad b\ r_N^>\ c$
- For $V = \{1, 2\}$, $\quad p \in V_{bc}$,
- $b\ q_V^>\ c$ but $q \notin V_{bc}$, and $a\ s_V^>\ c$ but $s \notin V_{ac}$

Social welfare functions

- A social welfare function is a function $F : \mathcal{D} \to O(X)$ assigning an ordering to each profile $p \in \mathcal{D}$.
- As above, $F(p)^>$ denotes the strict subrelation of $F(p)$.
- Fix a function $F : \mathcal{D} \to O(X)$. Define a relation $D_U \subseteq X \times X$ on the candidates for each $U \subseteq N$ as follows

$$a \ D_U \ b \text{ iff } a \neq b \text{ and for all } p \in U_{ab},\ a \ F(p)^> \ b$$

- Fix a function $F : \mathcal{D} \to O(X)$. Define a relation $E_U \subseteq X \times X$ on the candidates for each $U \subseteq N$ as follows

$$a \ E_U \ b \text{ iff for all } p \in \mathcal{D}, \text{ if } a \ p_U^> \ b, \text{ then } a \ F(p)^> \ b$$

Arrow's axioms

- **Universal Domain** For all $p \in L(\{a, b, c\})^n$, there exist $q \in \mathcal{D}$ such that $q|_{\{a,b,c\}} = p$
- **Weak Pareto** For all $p \in \mathcal{D}$, if $a \ p_N^> \ b$, then $a \ F(p) \ b$
- **Pareto** For all $p \in \mathcal{D}$, if $a \ p_N^> \ b$, then $a \ F(p)^> \ b$
- **Independence of Irrelevant Alternatives** For all $a, b \in X$, for all $p, q \in \mathcal{D}$, if $p|_{\{a,b\}} = q|_{\{a,b\}}$, then $F(p)|_{\{a,b\}} = F(q)|_{\{a,b\}}$

Arrovian dictator

- A voter $d \in N$ is a dictator if and only if for all profile $p \in \mathcal{D}$, for all candidates $a, b \in X$, if $a \ p_i^> \ b$, then $a \ F(p)^> \ b$.
- A voter $d \in N$ is a dictator iff for all $a, b \in X$, $a \ E_{\{d\}} \ b$.
- A social welfare function is a **dictatorship** provided there is a dictator.
- An example of a social welfare function that is a dictatorship (where voter i is a dictator) is $F_i(p) = p_i$ for all $p \in \mathcal{D}$. However, note that there are other functions that qualify as dictatorships. All that is required is that there is a voter d such that for any two candidates a, b if d ranks a above b, then society must rank a above b.

The theorem

Proposition 1 *For all $a, b, c \in X$,*

- *If $c \neq a$, then if $a\ D_U\ b$, then $a\ D_U\ c$*
- *If $c \neq b$, then if $a\ D_U\ b$, then $c\ D_U\ b$*

Lemma 2 *Suppose that R is an irreflexive relation on a set X with at least three elements such that, for all $a, b \in X$:*

1. *If $x \neq a$, then $a\ R\ b$ implies $a\ R\ x$, and*
2. *If $x \neq b$, then $a\ R\ b$ implies $x\ R\ b$.*

Then, if $x, y \in X$ are distinct, then $a\ R\ b$ implies $x\ R\ y$.

Proof. Suppose that R is an irreflexive relation on X and $a, b, x, y \in X$. Further, suppose that (1) and (2) hold. Suppose that $a\ R\ b$. Then, since R is irreflexive, $a \neq b$. We have three cases:

1. $y \neq a$: Then, $a\ R\ b$ implies $a\ R\ y$ (by 1.). Furthermore, $a\ R\ y$ implies $x\ R\ y$ (by 2. since $x \neq y$)
2. $x \neq b$: Then, $a\ R\ b$ implies $x\ R\ b$ (by 2.). Furthermore, $x\ R\ b$ implies $x\ R\ y$ (by 1. since $x \neq y$)
3. $y = a$ and $x = b$: Then, we must show $a\ R\ b$ implies $b\ R\ a$. Since X has at least three elements, there is a $c \in X$ such that $c \neq a$ and $c \neq b$. Then, $a\ R\ b$ implies $a\ R\ c$ (by 1. since $c \neq a$). Furthermore, $a\ R\ c$ implies $b\ R\ c$ (by 2. since $b \neq c$). Finally, $b\ R\ c$ implies $b\ R\ a$ (by 1. since $a \neq b$).

QED

Proposition 3 *For all $a, b, x, y \in X$ with $x \neq y$: if $a\ D_U\ b$ then $x\ D_U\ y$.*

Proof. This is an immediate consequence of Proposition 1 and Lemma 2. QED

Proposition 4 *For all $a, b \in X$, $a\ D_U\ b$ iff $a\ E_U\ b$*

The decisive sets: $\mathcal{U} = \{U \mid \text{ there are } a, b \in X \text{ such that } a\ D_U\ b\}$

Proposition 5 *The following are properties of* $\mathcal{U}$*:*

1. *For all* $U \subseteq N$*, either* $U \in \mathcal{U}$ *or* $U^C \in \mathcal{U}$
2. $N \in \mathcal{U}$
3. *For all* $U, V \subseteq N$*, if* $U \in \mathcal{U}$ *and* $U \subseteq V$*, then* $V \in \mathcal{U}$*.*
4. *For all* $U, V \in \mathcal{U}$*,* $U \cap V \in \mathcal{U}$*.*

Theorem 6 (Arrow's Theorem) *Assuming there are finitely man voters, at least three candidates, and all of Arrow's Axioms, there is a voter* $d \in N$ *such that* $\{d\} \in \mathcal{U}$*.*

Proof. We will show that there is some $d \in N$ such that $\{d\} \in \mathcal{U}$. Suppose that $N = \{1, 2, \ldots, n\}$. By Proposition 5 (2), we have $N \in \mathcal{U}$ (so $\mathcal{U}$ is nonempty). By Proposition 5 (1), we have

1. either $\{1\} \in \mathcal{U}$ or $\{2, 3, \ldots, n\} \in \mathcal{U}$. If $\{1\} \in \mathcal{U}$, then we are done (let $d = 1$). If not, then $\{2, 3, \ldots, n\} \in \mathcal{U}$.

2. either $\{2\} \in \mathcal{U}$ or $\{1, 3, \ldots, n\} \in \mathcal{U}$. If $\{2\} \in \mathcal{U}$, then we are done (let $d = 2$). If not, then $\{1, 3, \ldots, n\} \in \mathcal{U}$.

$\vdots$

$n-1$. either $\{n-1\} \in \mathcal{U}$ or $\{1, 2, \ldots, n-2, n\} \in \mathcal{U}$. If $\{n-1\} \in \mathcal{U}$, then we are done (let $d = n-1$). If not, then $\{1, 2, \ldots, n-2, n\} \in \mathcal{U}$.

If we have not found a dictator in any of the 1 to $n-1$ cases, then by proposition 5 (4),

$$\{n\} = \bigcap_{i=1}^{n-1} N - \{i\} = \{2, 3, \ldots, n\} \cap \{1, 3, \ldots, n\} \cap \cdots \cap \{1, 2, 3 \ldots, n-2, n\} \in \mathcal{U}$$

QED