Arrow's Theorem on Fair Elections

JWR

October 27, 2008

1 Introduction

The fair way to decide an election between two candidates a and b is majority rule; if more than half the electorate prefer a to b, then a is elected; otherwise b is elected. Arrow's theorem asserts that no fair election procedure[1] exists for choosing from among three or more candidates. This paper gives an exposition of Arrow's theorem. I learned the ultrafilter proof given below at a mathematics-economics conference held at the Univerisity of Warwick around 1975.

2 Informal examples

To get a feeling for Arrow's theorem let us consider how some existing election procedures can lead to grossly unfair results. (I downloaded much of this stuff from Wikipedia.)

One commonly used procedure is to have a second "runoff" election between the top two candidates if no candidate achieves a majority in the first election. The electorate might be confronted with three candidates a, b and c with candidates a and b extreme but opposite and c moderate. Suppose that each of the three candidates is the first choice of a third of the electorate and that all the supporters of a and b have c as their second choice. It seems clear that c is the best choice, especially if the supporters of a detest b and the supporters of b detest a. However, under the runoff procedure the electorate might well be forced to choose between a and b in the second election.

[1] Election procedures are sometimes called social choice functions.

Preferential voting (also called "instant runoff voting") is a type of ballot structure used in several electoral systems in which voters rank a list or group of candidates in order of preference. The candidate receiving the least first place votes is eliminated and his votes (with the preferences shifted up) are distributed among the remaining candidates. The process is repeated until only one candidate remains. Preferential voting is used in Australia, but the term "Australian Ballot" most commonly means simply "secret ballot".

Condercet voting works as follows: Rank the candidates in order (1st, 2nd, 3rd, etc.) of preference. Comparing each candidate on the ballot to every other, one at a time (pairwise), tally a "win" for the victor in each match. Sum these wins for all ballots cast. The candidate who has won every one of their pairwise contests is the most preferred, and hence the winner of the election.

3 Formulation of the Theorem

Throughout E denotes the **electorate**. The elements of this set represent the individual people who actually do the voting. For the formulation of the theorem we need assume nothing about this set except that it is nonempty and finite. The set of **candidates** is another finite set C. We assume only that the cardinality

$$n := |C|$$

of the set of candidates is at least 3. A **state** of the electorate is a function s which assigns to each individual elector $i \in E$ a simple ordering of the set C of candidates. A simple ordering of C is the same as a bijection between C and $\{1, 2, \ldots, n\}$ so a state of the electorate is a map $s : E \times C \to \{1, 2, \ldots, n\}$ such that for each $i \in E$ the function $s_i : C \to \{1, 2, \ldots, n\}$ defined by

$$s_i(a) = s(i, a)$$

is a bijection. The idea is that if the state of the electorate is s, then elector i prefers candidate a to candidate b if and only if $s(i, a) > s(i, b)$. We denote by Σ the set of all states of the electorate. The cardinality of Σ is

$$|\Sigma| = (n!)^{|E|}$$

where $|E|$ is the cardinality of E and $n = |C|$ is the cardinality of C. Evidentally, Σ is a rather large set.

An **election procedure** is a function which assigns to each state of the electorate an ordering of the candidates (the result of the election). In other words, an election procedure is a function

$$f : \Sigma \to \{1, 2, \ldots, n\}.$$

The condition $f(s)(a) > f(s)(b)$ says that the election procedure f ranks candidate a ahead of candidate b when the state of the electorate is s.

The election procedure f is said to satisfy the **unanimity condition** iff for all state s and all candidates a and b we have

$$\forall i \in E[s(i, a) > s(i, b)] \implies f(s)(a) > f(s)(b).$$

(If all electors favor Alice over Bob, then the election favors Alice over Bob.)

The election procedure f is said to satisfy the **monotonicity condition** iff whenever $s, s' \in \Sigma$ and $a, b \in C$ satisfy $f(s)(a) > f(s)(b)$, $f(s')(c) = f(s)(c)$ for all $c \in C \setminus \{a, b\}$, and $s_i(a) > s_i(b) \implies s'_i(a) > s'_i(b)$, then $f(s')(a) > f(s')(b)$. (If in a second run of an election in which the electorate favored Alice over Bob, some of the electors who previously voted for Bob over Alice now place Alice ahead of Bob but except for this no elector changes his/her vote, then Alice beats Bob in the second election as well.)

The election procedure f is said to satisfy the **irrelevance of third alternatives condition** iff whenever $s, s' \in \Sigma$ and $a, b \in C$ satisfy the condition that

$$s_i(a) > s_i(b) \iff s'_i(a) > s'_i(b) \text{ for all } i \in E$$

and $f(s)(a) > f(s)(b)$, then also $f(s')(a) > f(s')(b)$. (Whether or not the the election favors Alice over Bob has nothing to do with how the individual electors feel about Charles.

A **dictator** for election procedure f is an elector $i \in E$ whose preferences always coincide with the result of the election. In other words, $j \in E$ is a dictator for f iff for all states $s \in \Sigma$ we have $f(s) = s_j$. We would hardly call an election procedure fair if it has a dictator, but:

Arrow's Theorem: *Any election procedure which satisfies the unanimity condition, the monotonicity condition, and the irrelevance of third alternatives condition has a dictator.*

4 Ultrafilters

Let E be a nonempty set. A **filter** on E is a set $\mathcal{F}$ of subsets of E satisfying the following three conditions:

1. $E \in \mathcal{F}$ and $\emptyset \notin \mathcal{F}$;
2. If $X \subseteq Y \subseteq E$ and $X \in \mathcal{F}$, then $Y \in \mathcal{F}$;
3. If $X \in \mathcal{F}$ and $Y \in \mathcal{F}$, then $X \cap Y \in \mathcal{F}$.

Example 4.1. Let Z be any nonempty subset of E. Then the set

$$\mathcal{F} = \{X \subseteq E : Z \subseteq X\}$$

is a filter called the **principal filter** generated by Z.

Example 4.2. Let E be any infinite set. Then the set $\mathcal{F}$ of cofinite subsets of E is a filter. (A subset $\alpha \subseteq E$ is called **cofinite** iff its complement $E \setminus \alpha$ is finite.) This filter is not principal.

Theorem 4.3. *Let $\mathcal{F}$ be a filter on E. Then the following conditions are equivalent:*

1. *$\mathcal{F}$ is a maximal filter, i.e. if $\mathcal{F}'$ is a filter on E and $\mathcal{F} \subseteq \mathcal{F}'$ then $\mathcal{F} = \mathcal{F}'$;*
2. *$\mathcal{F}$ is a prime filter, i.e. if $X, Y \subseteq E$ and $X \cup Y \in \mathcal{F}$ then either $X \in \mathcal{F}$ or $Y \in \mathcal{F}$;*
3. *For every $X \subseteq E$ either $X \in \mathcal{F}$ or $E \setminus X in \mathcal{F}$;*
4. *If $Y \subseteq E$ and $Y \cap X \in \mathcal{F}$ for all $X \in \mathcal{F}$, then $Y \in \mathcal{F}$;*
5. *If $\{X_1, X_2, \ldots, X_r\}$ is a partition of E, then there is a j (necesarily unique) such that $X_j \in \mathcal{F}$.*

A filter which satisfies these equivalent conditions is called an **ultrafilter***.*

Example 4.4. A principal ultrafilter is a principal filter which is an ultrafilter. A principal filter is an ultrafilter if and only if its generator Z consists of a single point.

Theorem 4.5. *Every filter extends to an ultrafilter.*

Corollary 4.6. *There exist nonprincipal ultrafilters on an infinite set.*

Theorem 4.7. *On a finite set every filter is principal.*

5 Proof of Arrow's Theorem

Let f be an election procedure which satisfies the unanimity condition, the monotonicity condition, and irrelevance of third alternatives condition. For any state s and pair of candidates a and b define

$$P(s,a,b) = \{i \in E : s(i,a) > s(i,b)\}$$

denote the set of electors who prefer a to b. For $a \neq b$ call a set $X \subset E$ such that

$$\forall s \in \Sigma \left[X \subset P(s,a,b) \implies f(s)(a) > f(s)(b)\right]$$

a **forcing coalition for** a **over** b. Call the set X a **forcing coalition** iff for all $a \neq b$ it is a forcing coalition for a over b. Let $\mathcal{F}(a,b)$ be the set of forcing coalitions for a over b and

$$\mathcal{F} := \bigcap_{a \neq b} \mathcal{F}(a,b)$$

denote the set of all forcing coalitions.

Theorem 5.1. *Assume that the set C of candidates contains at least three members. Then $\mathcal{F}$ is an ultrafilter.*

Arrow's Theorem is an immediate corollary. By definition an elector $j \in E$ is a dictator iff the singleton $\{j\}$ is a forcing coalition. Since E is finite the ultrafilter $\mathcal{F}$ is principal and the generator is the dictator. The pattern of proof is as follows. First we show that each $\mathcal{F}(a,b)$ is a filter. Then we show that $\mathcal{F}(a,b)$ is independent of the choice of $a \neq b$, i.e. $\mathcal{F}(a,b) = \mathcal{F}(a',b')$ for $a \neq b$ and $a' \neq b'$. Thus $\mathcal{F}(a,b) = \mathcal{F}$. Finally we show that the minimal element of $\mathcal{F}$ (i.e. the intersection of all $X \in \mathcal{F}$) is a singleton.

The unanimity axiom says that the entire electorate is a forcing coalition, i.e. $E \in \mathcal{F}(a,b)$. It is immediate that any superset of a forcing coalition is again a forcing coalition, i.e. $X \in \mathcal{F}(a,b)$ and $X \subset Y \implies Y \in \mathcal{F}(a,b)$.

We show that the intersection of two forcing coalitions is a forcing coalition. Choose $X, Y \in \mathcal{F}(a,b)$. Choose $a, b \in C$ and $s \in \Sigma$ with $X \cap Y \subseteq P(s,a,b)$. Choose $c \neq a, b$ (as $n \geq 3$) and let $s' \in \Sigma$ satisfy $s'_i(a) > s'_i(c)$ for $i \in X \setminus Y$, $s'_i(a) > s'_i(c) > s'_i(b)$ for $i \in X \cap Y$, $s'_i(c) > s'_i(b)$ for $i \in Y \setminus X$, and $P(s',a,b) = P(s,a,b)$. Then $f(s')(a) > f(s')(c)$ as $X \in \mathcal{F}$ and $f(s')(c) > f(s')(a)$ as $Y \in \mathcal{F}$ so $f(s')(a) > f(s')(c) > f(s')(b)$. Hence

$f(s)(b) > f(s)(a)$ by the irrelevance of third alternatives. This proves $X \cap Y \in \mathcal{F}(a,b)$. We have shown that $\mathcal{F}(a,b)$ is a filter. Taking intersections shows that $\mathcal{F}$ is a filter.

Lemma 5.2. *For all states s and s' and all candidates a and b we have*

$$f(s)(a) > f(s)(b) \text{ and } P(s,a,b) \subseteq P(s',a,b) \implies f(s')(a) > f(s')(b).$$

Proof. Assume that $f(s)(a) > f(s)(b)$ and $P(s,a,b) \subseteq P(s',a,b)$. Define $s^* \in \Sigma$ by $s^* \in \Sigma$ by $s_i^*(a) = s_i'(a)$, $s_i^*(b) = s_i'(b)$ and $s_i^*(c) = s_i(c)$ for $i \in E$ and $c \neq a, b$. Then $f(s^*)(a) > f(s^*)(b)$ by monotonicity so $f(s')(a) > f(s')(b)$ by the irrelevance of third alternatives. □

Corollary 5.3. *Assume that for some $s \in \Sigma$ and $a, b \in C$ we have $f(s)(a) > f(s)(b)$ and let $X = P(s,a,b)$. Then $X \in \mathcal{F}(a,b)$.*

Lemma 5.4. $\mathcal{F}(a,b) = \mathcal{F}(a',b')$ *for all $a \neq b$ and $a' \neq b'$.*

Proof. Choose $X \in \mathcal{F}(a,b)$ and $a' \neq b$. We first show that $X \in \mathcal{F}(a',b)$. Assume that $X \subset P(s,a',b)$. Choose s' so that $P(s',a',b) = P(s,a,b)$. Then $f(s)(a) > f(s)(b)$ as $X \in \mathcal{F}(a,b)$ so $f(s')(a) > f(s')(b)$. This shows that $X \in \mathcal{F}(a',b)$. Hence $\mathcal{F}(a,b) \subset \mathcal{F}(a',b)$. Reversing the roles of a and a' gives $\mathcal{F}(a,b) = \mathcal{F}(a',b)$. Similarly $\mathcal{F}(a,b) = \mathcal{F}(a,b')$. Since a is arbitrary, $\mathcal{F}(a',b) = \mathcal{F}(a',b')$. Hence $\mathcal{F}(a,b) = \mathcal{F}(a',b) = \mathcal{F}(a',b')$ as required. □

We show that $\mathcal{F}$ is an ultrafilter. Let X be the minimal element of $\mathcal{F}$. Choose $j \in X$. We will show that $X = \{j\}$, i.e. that $X \setminus \{j\} = \emptyset$. Choose a, b, c distinct and a state $s \in \Sigma$ such that $s_j(a) > s_j(b) > s_j(c)$, $s_i(c) > s_i(a) > s_i(b)$ for $i \in X \setminus \{j\}$, and $s_i(b) > s_i(c) > s_i(a)$ for $i \in E \setminus X$. Since $P(s,a,b) = X$, we have $f(s)(a) > f(s)(b)$. Now $P(s,c,b) = X \setminus \{j\}$. If it were the case that $f(s)(c) > f(s)(b)$, then (by Corollary 5.3) $X \setminus \{j\}$ would force c over b, so $X \setminus \{j\} \in \mathcal{F}(c,b)$, so (by Lemma 5.4) $X \setminus \{j\} \in \mathcal{F}$ contradicting minimality. Hence $f(s)(b) > f(s)(c)$. It follows that $f(s)(a) > f(s)(c)$. But $P(s,a,c) = \{j\}$. Hence, by Corollary 5.3 we have $\{j\} \in \mathcal{F}(a,c) = \mathcal{F}$. This completes the proof.

6 Reflections

When I have discussed Arrow's theorem with nonmathematicians I discover that they tend to attack the theorem by attacking its assumptions. This is of course quite reasonable, but the nonmathematicians do this by trying to impose additional assumptions. They say something like "Well of course you reached an antidemocratic solution: your hypotheses didn't assume all members of the electorate are equal!" What they don't understand is that additional hypotheses cannot possibly falsify a true theorem.

It is tempting to conclude that the theorem proves something about political life like the most stable countries are those which have a two party system. Possibly some people might even take the theorem as an argument against democracy. I am skeptical of such inferences. It seems to me that democracy is successful when all voices are heard and the citizenry understand one another and have some control over their fate. I don't see what Arrow's theorem says about that.