Fair Elections

JWR

November 30, 2012

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 exists for choosing from among three or more candidates. This note gives an exposition of Arrow's theorem. It also describes the relation of election procedures between two candidates to Dedekind numbers.

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.

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".) This process can produce a bad result in the same way that a second runoff election does.

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 winner of the election. But if there are three candidates a, b, c and three electors with preferences $a > b > c$, $c > a > b$, $b > c > a$ then a defeats b, b defeats c, and c defeats a so the procedure does not produce an outcome.

Even when there are only two candidates paradoxical results can occur. For example if we have an election where b wins over a and then a rerun where everyone who voted for b over a does so again, it seems fair that b should win over a in the rerun as well. This is called the *monotonicity rule* below. But consider an even split of the US Senate. The vice president might prefer a to b, but if one senator changes his/her vote from a over b to b over a, then the outcome changes from a over b to b over a and the monotonicity rule is violated. The fact that the vice president usually has no vote doesn't destroy this example; we could imagine an amendment to the US constitution such that there is a single senator from the District of Columbia and the vice president has full voting rights and in addition preserves the power to break ties. The same paradox can occur.

3 Definitions

Our objective is to give a precise definition for what is a "fair election procedure". First we define what is an "election procedure".

§1. Setup. Throughout E denotes a nonempty set called the **electorate** and C denotes a nonempty set called the set of **candidates**. The elements of the set E represent the individual people who actually do the voting. A **state** of the electorate is a function which assigns to each individual elector $x \in E$ a linear ordering of the set C of candidates. Denote the set of linear orderings of C by $O(C)$ so a state is a map $\sigma : E \to O(C)$ and

$$\Sigma := O(C)^E$$

denotes the set of all states. The set Σ has cardinality $|\Sigma| = (n!)^{|E|}$ where $n = |C|$ is the cardinality of C and $|E|$ is the cardinality of E.

§ 2. Definition. An **election procedure** is a function which assigns to each state of the electorate an ordering of the candidates (the outcome of the election).[1] In other words, an election procedure is a map

$$f : \Sigma \to O(C).$$

§3. Remark. The above definition is open to the criticism that it does not correctly model the problem of constructing a "fair election procedure" in that (i) it forces every elector to order his/her preferences linearly whereas electors might not distinguish between candidates they don't like, and (ii) it forces an outcome which is a linear order even though all that is required in a real election is a single winner. Such criticisms are discussed in [8].

§4. Notation. Introduce the abbreviation

$$C_2 = \{(a, b) \in C^2 : a \neq b\}$$

for the set of ordered pairs of distinct elements of C. For $\lambda \in O(C)$ and $a, b \in C$ write

$$a <_\lambda b$$

to mean that a precedes b in the order λ. Thus for $a, b \in C$ the notation $a <_{\sigma(x)} b$ means that elector x prefers candidate b to candidate a when the state of the electorate is σ and the notation $a <_{f(\sigma)} b$ means that the election procedure f ranks candidate b ahead of candidate a when the state of the electorate is σ. For $(a, b) \in C_2$ and $\sigma \in \Sigma$ let

$$P_{ab}(\sigma) := \{x \in E : \ a <_{\sigma(x)} b\}$$

denote the set of electors x who prefer b to a when the state is σ. The conditions

$$E = P_{ab}(\sigma) \cup P_{ba}(\sigma), \quad P_{ab}(\sigma) \cap P_{ba}(\sigma) = \emptyset, \quad P_{ab}(\sigma) \cap P_{bc}(\sigma) \subseteq P_{ac}(\sigma)$$

for $(a, b) \in C_2$ express the condition that $<_{\sigma(x)}$ is a linear order for $x \in E$.

---

[1] I have seen other terminology in the literature. For example, in [8] an election procedure is called a *social welfare function* and a state is called a *profile of the electorate*.

§5. A **linear order** on a set C is a relation (i.e. a subset of C^2) satisfying

$$\text{(Irreflexivity)} \qquad a < b \iff \neg b < a$$

for $(a, b) \in C^2$ and

$$\text{(Transitivity)} \qquad a < b \text{ and } b < c \implies a < c.$$

Since these properties are preceded by an implicit universal quantifier on a, b, c the empty relation is always an example. The irreflexivity condition implies that $a < b \implies a \neq b$. If $|C| = 1$ then $C_2 = \emptyset$ and $O(C) = \{\emptyset\}$. If $|C| = 2$ then $|C_2| = 2$ and the transitivity condition holds vacuously so $|O(C)| = 2$.

§6. The election procedure f satisfies the **strong unanimity condition** iff

$$P_{ab}(\sigma) = E \implies a <_{f(\sigma)} b \tag{SU}$$

for all $(a, b) \in C_2$. The election procedure f satisfies the **weak unanimity condition** iff

$$\sigma^{-1}(\lambda) = E \implies f(\sigma) = \lambda \tag{WU}$$

for all $\lambda \in O(C)$.

§7. Remark. The weak unanimity condition (WU) implies that $f^{-1}(\lambda) \neq \emptyset$ for every $\lambda \in O(C)$. This in turn implies that for every pair $(a, b) \in C_2$ there exists $\sigma \in \Sigma$ such that $a <_{f(\sigma)} b$ This last condition is called *citizen's sovereignty* in [8].

§8. The election procedure f satisfies the **monotonicity condition** iff

$$\left.\begin{array}{ll} P_{cd}(\sigma) = P_{cd}(\tau) & \forall c, d \in C \setminus \{b\}, \\ P_{cb}(\sigma) \subseteq P_{cb}(\tau) & \forall c \in C \setminus \{b\}, \text{ and} \\ a <_{f(\sigma)} b & \end{array}\right\} \implies a <_{f(\tau)} b. \tag{M}$$

In other words

> *If in a second run of an election in which the electorate favored b over a, some of the electors change their preferences between b and other candidates in favor of b but except for this no elector changes his/her vote, then b finishes ahead of a in the second election as well.*

§9. The election procedure f satisfies the **independence condition** iff

$$P_{ab}(\sigma) = P_{ab}(\tau) \;\blacksquare\!\implies\!\blacksquare\; a <_{f(\sigma)} b \iff a <_{f(\tau)} b.$$

In other words

> *Whether or not an election favors b over a is independent of how the individual electors feel about the other candidates.*

§10. Remark. In [8] pages 331-340 an election procedure is called a *social welfare function*, a state is called a *profile*, the monotonicity condition §8(M) is called the *positive association of social and individual values*, and the independence condition §9(I) is called the *independence of irrelevant alternatives*.

§11. If $|C| = 1$ then $C_2 = \emptyset$ and $O(C) = \{\emptyset\}$ so there is only one election procedure, namely $f(x) = \emptyset$ for $x \in X$, i.e. the lone candidate always wins (since s/he finishes ahead of all the other candidates) and always loses (since s/he finishes behind all the other candidates).

If $|C| = 2$ then $|C_2| = 2$ and the hypothesis $P_{ab}(\sigma) = P_{ab}(\tau)$ of the independence condition (I) of §9 implies $\sigma = \tau$ so (I) automatically holds, however the monotonicity condition (M) of§8 has some content. For example, if $|E|$ is odd, majority rule satisfies unanimity and monotonicity. See Appendix 6.

If $|C| = 2$, weak and strong unanimity are equivalent. If there are three or more candidates, condition (WU) is weaker than (SU) as there are states where the electorate unanimously prefers one candidate to another but is not unanimous on the other candidates. However, if f satisfies the independence condition (I), then the following lemma says that (WU) and (SU) are equivalent.

§12. Lemma. *An election procedure which satisfies the strong unanimity condition (SU) of §6 also satisfies the weak unanimity condition (WU) of §6. If the election procedure satisfies the independence condition (I) from (§9), then the converse holds.*

Proof. Assume (SU) and $\sigma^{-1}(\lambda) = E$. (We must show that $f(\sigma) = \lambda$.) As $\sigma^{-1}(\lambda) = \bigcap_{a<_\lambda b} P_{ab}(\sigma)$ it follows that $P_{ab}(\sigma) = E$ for all $(a, b) \in C_2$ such that $a <_\lambda b$. Hence from (SU) it follows that $a <_\lambda b \implies a <_{f(\sigma)} b$. For $\lambda, \mu \in O(C)$ the condition $a <_\lambda b \implies a <_\mu b$ for all $(a, b) \in C_2$ holds if and only if $\lambda = \mu$. (If $\lambda \neq \mu$ then $a <_\lambda b$ but $b <_\mu a$ for some $(a, b) \in C_2$.) Hence $f(\sigma) = \lambda$. This proves (SU) $\implies$ (WU).

Conversely assume (WU) and choose $(a, b) \in C_2$ satisfying $P_{ab}(\sigma) = E$. (We must show that $a <_{f(\sigma)} b$.) Choose an enumeration $C = \{c_1, \ldots, c_n\}$ with $a = c_1$ and $b = c_2$ and define $\lambda \in O(C)$ by $c_i <_\lambda c_j \iff i < j$. Let $\tau : E \to O(C)$ be the constant map $\tau(x) = \lambda$. Then $P_{ab}(\tau) = E$ so $P_{ab}(\tau) = P_{ab}(\sigma)$. By (I) $a <_{f(\sigma)} b \iff a <_{f(\tau)} b$ and by (WU) $a <_{f(\tau)} b$. Hence $a <_{f(\sigma)} b$. Thus proves (WU) $\implies$ (SU). □

§ 13. Dictator. A **dictator** for election procedure f is an elector $z \in E$ whose preferences always coincide with the result of the election. In other words, $z \in E$ is a dictator for f iff for all states $\sigma \in \Sigma$ we have $f(\sigma) = \sigma(z)$, i.e.

$$a <_{f(\sigma)} b \iff a <_{\sigma(z)} b \qquad (D)$$

for all $(a, b) \in C_2$. It is immediate that an election procedure which has a dictator satisfies the other conditions (WU), (SU), (M), and (I). We would hardly call an election procedure fair if it has a dictator but

§14. Arrow's Theorem (finite version). *Let $f : E \to O(C)$ be an election procedure with a nonempty finite electorate E and a finite set C of at least three candidates. Then f satisfies the weak unanimity condition (WU) of §6 and independence condition (I) of §9 conditions if and only if there is a dictator z for f.*

See http://en.wikipedia.org/wiki/Arrow's_impossibility_theorem for an informal proof. There is a proof in [8] and a citation to the original paper of Arrow appears there. In §23 we will prove a generalization which includes the case where the electorate E is infinite. Note that the monotonicity condition §8(M) is not assumed in Arrow's Theorem.

4 Some Preliminary Lemmas

In this section $X \subseteq E$ is a fixed subset of the electorate and $f : E \to O(C)$ denotes an election procedure satisfying the weak unanimity condition (WU) of §6, the independence condition (I) of §9, and hence (by the lemma of §12) the strong unanimity condition (SU) of §6.

§15. Lemma. *For $X \subseteq E$ and $(a, b) \in C_2$ the following are equivalent.*

(i) $\exists \sigma_0 \in \Sigma$ ▪ $X = P_{ab}(\sigma_0)$ *and* $a <_{f(\sigma_0)} b$.

(ii) $\forall \sigma \in \Sigma$ ▪ $X = P_{ab}(\sigma) \implies a <_{f(\sigma)} b$.

Proof. To prove (ii) $\implies$ (i) assume (ii). (We must construct σ_0 satisfying the two conditions $X = P_{ab}(\sigma_0)$ and $a <_{f(\sigma_0)} b$.) Define $\sigma_0 : E \to O(\{a, b\})$

$$\begin{array}{ll} a <_{\sigma_0(x)} b & \text{if } x \in X, \\ b <_{\sigma_0(x)} a & \text{otherwise.} \end{array}$$

For each $x \in X$ extend the order $\sigma_0(x)$ to an element of $O(C)$ arbitrarily. Then $\sigma_0 \in \Sigma$ and $X = P_{ab}(\sigma_0)$. Read σ_0 for σ in (ii) to conclude $a <_{f(\sigma_0)} b$. This proves (ii) $\implies$ (i).

To prove that (i) $\implies$ (ii) assume (i) and choose σ such that $X = P_{ab}(\sigma)$. (We must show $a <_{f(\sigma)} b$.) By (i) $X = P_{ab}(\sigma_0)$ so $P_{ab}(\sigma) = P_{ab}(\sigma_0)$. Therefore $a <_{f(\sigma)} b \iff a <_{f(\sigma_0)} b$ by the independence condition (§9). But $a <_{f(\sigma_0)} b$ by (i) so $a <_{f(\sigma)} b$. This proves (i) $\implies$ (ii). □

§16. Lemma. *Fix a subset $X \subseteq E$ and for each $(a, b) \in C_2$ introduce the abbreviation*[2]

$$D(a, b) \leftrightsquigarrow \forall \sigma \in \Sigma \;▪\; X = P_{ab}(\sigma) \implies a <_{f(\sigma)} b.$$

Then for all distinct $a, b, c \in C$ we have

$$D(a, b) \implies D(c, b) \qquad \text{and} \qquad D(a, b) \implies D(a, c).$$

Proof. To prove that $D(a, b) \implies D(c, b)$ assume $D(a, b)$ and choose $\sigma \in \Sigma$ satisfying $X = P_{cb}(\sigma)$. (We must show $c <_{f(\sigma)} b$.) Define $\tau : E \to O(\{a, b, c\})$ by

$$\begin{array}{ll} c <_{\tau(x)} a <_{\tau(x)} b & \text{if } c <_{\sigma(x)} b, \\ b <_{\tau(x)} c <_{\tau(x)} a & \text{if } b <_{\sigma(x)} c, \end{array}$$

and extend each order $\tau(x)$ to an element of $O(C)$ arbitrarily. Then

$$P_{ca}(\tau) = E \quad \text{and} \quad P_{ab}(\tau) = P_{cb}(\tau) = P_{cb}(\sigma) = X.$$

Now $c <_{f(\tau)} a$ by unanimity (§6) and $a <_{f(\tau)} b$ by reading τ for σ in (ii). Hence $c <_{f(\tau)} b$ as $<_{f(\tau)}$ is transitive. But $P_{cb}(\tau) = P_{cb}(\sigma)$ so $c <_{f(\sigma)} b$ by independence (§9). This proves $D(a, b) \implies D(c, b)$.

[2] The notation $\leftrightsquigarrow$ means that the formula on the left is an abbreviation for the for the one on the right.

The proof that $D(a,b) \implies D(a,c)$ is similar. Assume $D(a,b)$ and choose $\sigma \in \Sigma$ satisfying $X = P_{ac}(\sigma)$. (We must show that $a <_{f(\sigma)} c$.) As in the proof of $D(a,b) \implies D(c,b)$ construct τ so that

$$P_{bc}(\tau) = E \quad \text{and} \quad P_{ab}(\tau) = P_{ac}(\tau) = P_{ac}(\sigma) = X.$$

Then $b <_{f(\tau)} c$ by unanimity and $a <_{f(\tau)} b$ by $D(a,b)$, so $a <_{f(\tau)} c$ by transitivity and hence $a <_{f(\sigma)} c$ by independence. Any τ satisfying

$$\begin{array}{ll} a <_{\tau(x)} b <_{\tau(x)} c & \text{if } a <_{\sigma(x)} c, \\ b <_{\tau(x)} c <_{\tau(x)} a & \text{if } c <_{\sigma(x)} a \end{array}$$

has the desired properties. □

§ 17. Corollary. Assume that $|C| \geq 3$. Then if $D(a,b)$ holds for some $(a,b) \in C_2$ it holds for all $(a,b) \in C_2$.

Proof. Assume $D(a,b)$. Choose $c \in C$ so that a, b, c are distinct. Then $D(c,b)$ and $D(a,c)$ by §16. Read (c,b) for (a,b) in §16 to conclude $D(c,a)$ and read (a,c) for (a,b) in §16 to conclude $D(b,c)$. Finally $D(c,a) \implies D(b,a)$ follows from §16 by reading (b,c,a) for (a,b,c) so $D(b,a)$ holds as we have already proved $D(c,a)$. □

5 Proof of Arrow's Theorem

Throughout this section $f : E \to O(C)$ denotes an election procedure satisfying the weak unanimity condition (WU) of §6, the independence condition (I) of (§9), and hence (by the lemma of §12) the strong unanimity condition (SU) of §6. We also assume that $|C| \geq 3$, i.e. that there are at least three candidates. For $X \subseteq E$ let

$$X' := E \setminus X$$

denote the complement of X in E.

§18. Definition. A subset $X \subseteq E$ is called a **decisive** iff

$$X \subseteq P_{ab}(\sigma) \implies a <_{f(\sigma)} b$$

for all $(a,b) \in C_2$. The set of decisive subsets will be denoted by $\mathcal{D}$.

§19. $\emptyset \notin \mathcal{D}$.

Proof. The set Σ is nonempty as E and C are and C_2 is nonempty as $|C| \geq 2$. Hence there exists at least one $\sigma \in \Sigma$ and and at least one pair $(a, b) \in C_2$ i.e. $a \neq b$ and at least two distinct elements λ and μ in $O(C)$, one where $a <_\lambda b$ and another where $b <_\mu a$. Now $\emptyset \subseteq \sigma^{-1}(\lambda)$ so $f(\sigma) = \lambda$ and $\emptyset \subseteq \sigma^{-1}(\mu)$ so $f(\sigma) = \mu$. This is a contradiction so $\emptyset \notin \mathcal{D}$. □

§20. $E \in \mathcal{D}$.

Proof. This is a reformulation of the unanimity condition (SU) of §6. □

§21. $X \in \mathcal{D}$ and $X \subseteq Y \implies Y \in \mathcal{D}$.

Proof. In other words, enlarging a decisive set gives another decisive set. To prove this assume $X \in \mathcal{D}$ and $X \subseteq Y$. To prove $Y \in \mathcal{D}$ choose $a, b \in C$, $\sigma \in \Sigma$, and assume $Y \subseteq P_{ab}(\sigma)$. (We must show $a <_{f(\sigma)} b$.) Then $X \subseteq P_{ab}(\sigma)$ as $X \subseteq Y$. Hence $a <_{f(\sigma)} b$ as $X \in \mathcal{D}$. As a, b, and σ were arbitrary this proves $Y \in \mathcal{D}$. □

§22. $X, Y \in \mathcal{D} \implies X \cap Y \in \mathcal{D}$.

Proof. In other words, the intersection of two decisive sets is again decisive. Choose $X, Y \in \mathcal{D}$. To show $X \cap Y \in \mathcal{D}$ choose $a, b \in C$ and $\sigma \in \Sigma$ with $X \cap Y \subseteq P_{ab}(\sigma)$; we must show that $a <_{f(\sigma)} b$. To do this it is enough to choose $c \neq a, b$ (this is possible as $|C| \geq 3$) and construct $\tau \in \Sigma$ satisfying

$$X \subseteq P_{ac}(\tau), \qquad Y \subseteq P_{cb}(\tau), \qquad P_{ab}(\tau) = P_{ab}(\sigma),$$

for then $a <_{f(\tau)} c$ (because $X \in \mathcal{D}$), $c <_{f(\tau)} b$ (because $Y \in \mathcal{D}$), $a <_{f(\tau)} b$ (by transitivity), and hence $a <_{f(\sigma)} b$ (by independence).

Abbreviate $P_{ab}(\sigma)$ by P so $P' = P_{ba}(\sigma)$. The following table defines a map $\tau : E \to O\big(\{a, b, c\}\big)$.

$P \cap X \cap Y$	$a < c < b$	$P' \cap X \cap Y$	$\emptyset$
$P \cap X \cap Y'$	$a < c < b$	$P' \cap X \cap Y'$	$b < a < c$
$P \cap X' \cap Y$	$a < c < b$	$P' \cap X' \cap Y$	$c < b < a$
$P \cap X' \cap Y'$	$a < b$	$P' \cap X' \cap Y'$	$b < a$

In the bottom row it does not matter where c fits into the order and the upper right hand entry is $\emptyset$ as $X \cap Y \subseteq P$, i.e. $P' \cap X \cap Y = \emptyset$. Now define τ to take values in $O(C)$ by extending each $\tau(x) \in O\big(\{a, b, c\}\big)$ to $O(C)$ arbitrarily. The state $\tau \in \Sigma$ has the desired properties. □

§23. Arrow's Theorem (Ultrafilter Version.) *$\mathcal{D}$ is an ultrafilter.*

Proof. §19, §20, §21, §22 say that $\mathcal{D}$ is a filter. To show that $\mathcal{D}$ is an ultrafilter we must show (see the theorem in §33) that for every subset $X \subseteq E$ either it or its complement $X' = E \setminus X$ is an element of $\mathcal{D}$. Choose $\sigma_0 \in \Sigma$ (possible as $\Sigma \neq \emptyset$) and distinct $a_0, b_0 \in C$ (possible as $|C| \geq 3$). Interchanging a and b if necessary we may assume that $a <_{f(\sigma_0)} b$. Let $X = P_{a_0 b_0}(\sigma_0)$. Then $D(a_0, b_0)$ holds by the lemma of §15. Hence by the corollary in §17 $D(a, b)$ holds for all $a, b \in C_2$, i.e. X is decisive, i.e. $X \in \mathcal{D}$. □

§24. I learned the ultrafilter formulation given below from a talk by Alan Kirman at a mathematics-economics conference held at the University of Warwick around 1975. It is due to Kirman and Sondermann (see [6]). See [7] for another exposition. The finite version of Arrow's Theorem (§14) is an immediate corollary. By definition, an elector $z \in E$ is a dictator iff the singleton $\{z\}$ is a decisive set. When E is finite, the ultrafilter $\mathcal{D}$ is principal by the theorem in §37 and the generator is the dictator.

6 Election Procedures With Two Candidates

In this section we consider election procedures between two candidates. In this case the independence axiom from §9 is automatic (there are no "irrelevant third alternatives") and there is no difference between the strong and weak unanimity conditions from §6. Any nonconstant election procedure with a finite electorate which satisfies the monotonicity condition (M) from §8 also satisfies the unanimity condition.

§25. Represent the two linear orderings of the two element set $C = \{a, b\}$ by 0 and 1, say $a <_0 b$ and $b <_1 a$. The set Σ of states of the electorate is the same as the power set of the set E and an election procedure $f : \Sigma \to \{0, 1\}$ satisfies the monotonicity condition (M) if and only if it is monotonic i.e.

$$X \subseteq Y \implies f(X) \leq f(Y)$$

for $X, Y \subseteq E$. (If $X = \sigma^{-1}(1)$ and $Y = \tau^{-1}(1)$ then the condition $X \subseteq Y$ is the same as the condition $\forall x \in E\,_{\blacksquare}\, \sigma(x) \leq \tau(x)$.) Let

$$E = \{x_1, \ldots, x_n\}$$

and for each $i = 1, \ldots, n$ let $e_i : \Sigma \to \{0, 1\}$ denote the election procedure which as x_i as dictator, i.e.

$$e_i(X) = \begin{cases} 1 & \text{if } x_i \in X, \\ 0 & \text{otherwise} \end{cases}$$

§26. The set of all election procedures $f : \Sigma \to \{0, 1\}$ forms a distributive lattice under the operations

$$(f \wedge g)(x) = \min\{f(x), g(x)\}, \qquad (f \vee g)(x) = \max\{f(x), g(x)\}.$$

The election procedures $e_1, \ldots, e_n$ generate this lattice, i.e. any election procedure f is expressible in the form

$$f = e_{I_1} \vee \cdots \vee e_{I_r}$$

for some collection $I_1, \ldots, I_r$ of subsets of $\{1, \ldots, n\}$ where

$$e_I := e_{i_1} \wedge \cdots \wedge e_{i_k} \text{ for } I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}.$$

§27. Each antichain $\{X_1, \ldots, X_r\}$ of subsets of E determines an election procedure f by

$$f(X) = \begin{cases} 1 & \text{if } X_i \subseteq X \text{ for some } i, \\ 0 & \text{otherwise} \end{cases}$$

(An antichain of subsets of E sets is a family of sets none of which is contained in any other set.) This defines a bijective correspondence between election procedures and antichains.

§28. Each election procedure $f : \Sigma \to \{0, 1\}$ determines an abstract simplicial complex with n vertices. The corresponding antichain is the set of maximal faces. This defines a bijective correspondence between the set of election procedures on electorate of size n and the set of abstract simplicial complexes with n vertices.

§29. Dedekind's Problem[3] [4] is to compute the number of monotonic election procedures as a function of the cardinality of E. Apparently this is quite difficult. According to [3] page 121 the first eight values are

[3]See http://en.wikipedia.org/wiki/Dedekind_number

1	3
2	6
3	20
4	168
5	7581
6	7828354
7	2414682040998
8	56130437228687557907788

§ 30. Each map $v : E \to \mathbb{R}^+$ determines an election procedure f via the formula

$$a <_{f(\sigma)} b \iff \sum_{b <_{\sigma(x)} a} v(x) < \sum_{a <_{\sigma(x)} b} v(x).$$

If v is such that the sums $\sum_{x \in X} v(x)$ are distinct, then this election procedure satisfies the unanimity condition (SU) of §6 and the monotonicity condition (M) of §8.

Here is an example[4] which shows that not every election procedure satisfying (SU) and (M) arises this way.

$P_{ab}(\sigma)$	$a <_{f(\sigma)} b$?	$\sum_{b <_{\sigma(x)} a} v(x) < \sum_{a <_{\sigma(x)} b} v(x)$
$\emptyset$	no	$0 < v_1 + v_2 + v_3 + v_4$
$\{x_1\}$	no	$v_1 < v_2 + v_3 + v_4$
$\{x_2\}$	no	$v_2 < v_1 + v_3 + v_4$
$\{x_3\}$	no	$v_3 < v_1 + v_2 + v_4$
$\{x_4\}$	no	$v_4 < v_1 + v_2 + v_3$
$\{x_1, x_2\}$	yes	$v_3 + v_4 < v_1 + v_2$
$\{x_1, x_3\}$	no	$v_1 + v_3 < v_2 + v_4$
$\{x_1, x_4\}$	no	$v_1 + v_4 < v_2 + v_3$
$\{x_2, x_3\}$	no	$v_2 + v_3 < v_1 + v_4$
$\{x_2, x_4\}$	no	$v_2 + v_4 < v_1 + v_3$
$\{x_3, x_4\}$	yes	$v_1 + v_2 < v_3 + v_4$
$\{x_1, x_2, x_3\}$	yes	$v_4 < v_1 + v_2 + v_3$
$\{x_1, x_2, x_4\}$	yes	$v_3 < v_1 + v_2 + v_4$
$\{x_1, x_3, x_4\}$	yes	$v_2 < v_1 + v_3 + v_4$
$\{x_2, x_3, x_4\}$	yes	$v_1 < v_2 + v_3 + v_4$
$\{x_1, x_2, x_3, x_4\}$	yes	$0 < v_1 + v_2 + v_3 + v_4$

[4]Eric Bach helped me with this. The example comes from [9].

7 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.

Appendices

A Ultrafilters

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

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

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

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

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

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

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

1. *$\mathcal{D}$ is a maximal filter, i.e. if $\mathcal{D}'$ is a filter on E and $\mathcal{D} \subseteq \mathcal{D}'$ then $\mathcal{D} = \mathcal{D}'$.*

2. *$\mathcal{D}$ is a prime filter, i.e. if $X, Y \subseteq E$ and $X \cup Y \in \mathcal{D}$ then either $X \in \mathcal{D}$ or $Y \in \mathcal{D}$.*

3. *For every $X \subseteq E$ either $X \in \mathcal{D}$ or $E \setminus X \in \mathcal{D}$.*

4. *For any partition $E = E_1 \cup \cdots \cup E_r$ of E into pairwise disjoint sets, $E_i \in \mathcal{D}$ for some (necessarily unique) i.*

5. *If $Y \subseteq E$ and $Y \cap X \in \mathcal{D}$ for all $X \in \mathcal{D}$, then $Y \in \mathcal{D}$.*

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

§34. Example. 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.

§35. Theorem. *Every filter extends to an ultrafilter.*

§36. Corollary. *There exist nonprincipal ultrafilters on an infinite set.*

§37. Theorem. *On a finite set every filter is principal.*

§38. Corollary. Every ultrafilter $\mathcal{D}$ on a finite set E has form

$$\mathcal{D} = \{X \subseteq E : z \in X\}$$

for some element $d \in E$.

B Birkhoff Representation

§ 39. A **poset** is a set P equipped with a partial order, i.e. a relation satisfying

(reflexivity) $x \leq x$,

(antisymmetry) $x \leq y$ and $y \leq x$ imply $x = y$,

(transitivity) $x \leq y$ and $y \leq z$ imply $x \leq z$,

for $a, b, c \in P$. The category of partially order sets and order preserving maps will be denoted by $\mathfrak{P}$, i.e. for posets $P, Q \in \mathfrak{P}$ the set of morphisms from P to Q is the set

$$\mathfrak{P}(P, Q) := \{\phi : R \to Q \ \ x \leq y \implies \phi(x) \leq \phi(y) \quad \forall x, y \in P\}.$$

§40. A **lattice** is a set L equipped with two binary operations $\wedge$ (**meet**) and $\vee$ (**join**) satisfying

(associativity) $a \vee (b \vee c) = (a \vee b) \vee c$ and $a \wedge (b \wedge c) = (a \wedge b) \wedge c$,

(commutativity) $a \vee b = b \vee a$ and $a \wedge b = b \wedge a$,

(idempotence) $a \vee a = a$ and $a \wedge a = a$,

(absorption) $a \vee (a \wedge b) = a$ and $a \wedge (a \vee b) = a$,

for all $a, b, c \in L$. Let $\mathfrak{L}$ denote the category of lattices and lattice homomorphisms, i.e. an object of $\mathfrak{L}$ is a lattice L for lattices $L, K \in \mathfrak{L}$ the set of morphisms from L to K is the set

$$\mathfrak{L}(L, K) := \left\{ f : L \to K, \quad \begin{array}{ll} f(a \wedge b) = f(a) \wedge f(b) & \text{and} \\ f(a \vee b) = f(a) \vee f(b) & \forall a, b \in L \end{array} \right\}.$$

§41. (Equivalent Definition). (i) For any lattice L we have

$$a \wedge b = a \iff a \vee b = b.$$

(ii) The relation defined by $a \leq b \iff a \wedge b = a$ is a partial order. It is the same as the relation defined by $a \leq b \iff a \vee b = b$.

(iii) A poset P arises from a lattice L in this way if and only if for all $a, b \in P$ the set $\{c \in P : a \leq c \text{ and } b \leq c\}$ has a unique minimal element (namely $a \wedge b$) and the set $\{c \in P : c \leq a \text{ and } c \leq b\}$ has a unique maximal element (namely $a \vee b$).

Proof. See [3] Theorem 2.10 page 40. □

§42. Lemma/Definition. A lattice L satisfies the identity

$$(a \vee b) \wedge c = (a \wedge c) \vee (a \wedge c)$$

for all $a, b, c \in L$ if and only if its satisfies the identity

$$(a \wedge b) \vee c = (a \vee c) \wedge (a \vee c).$$

A lattice which satisfies these equivalent identities is called **distributive**. We denote by $\mathfrak{D} \subset \mathfrak{L}$ the subcategory of distributive lattices.

Proof. See [3] Lemma 4.3 page 85. □

§43. By the definitions in §41 and §42 we have an inclusion of categories

$$\mathfrak{D} \subseteq \mathfrak{L} \subseteq \mathfrak{P}.$$

The category $\mathfrak{D}$ is a full subcategory of the category $\mathfrak{L}$, i.e. $\mathfrak{D}(L, K) = \mathfrak{L}(L, K)$ for $K, L \in \mathfrak{D}$, but the category $\mathfrak{L}$ is not a full subcategory of the category $\mathfrak{P}$. For example, subset of a lattice need not be closed under meet and join.

§44. Every poset P determines a **dual lattice**

$$P^* := \mathfrak{L}(P, \{0, 1\}).$$

The lattice operations are defined by

$$(\alpha \wedge \beta)(x) = \min(\alpha(x), \beta(x)), \qquad (\alpha \vee \beta)(x) = \max(\alpha(x), \beta(x)),$$

so $\alpha \leq \beta \iff \alpha(x) \leq \beta(x)$ for all $x \in P$. Each morphism $\phi \in \mathfrak{P}(P, Q)$ induces a lattice homomorphism $\phi^* \in \mathfrak{L}(Q^*, P^*)$ defined by

$$\phi^* \alpha = \alpha \circ \phi.$$

Thus $P \mapsto P^*$ is a contravariant functor from $\mathfrak{P}$ to $\mathfrak{L}$. In [3] a set of form $\sigma^{-1}(0)$ where $\sigma \in P^*$ is called a *down-set* of the poset P and a set of form $\sigma^{-1}(1)$ is *up-set* of P.[5]

[5] Commonly used alternate terminology for down-set: lower set, decreasing set, order ideal; for up-set: upper set, increasing set, order filter.

§ 45. Example. Let E be a set and equip it with the partial order of equality (so that any two distinct elements of E are incomparable). The power set of E is a lattice with the meet operation intersection and join operation union The rule which assigns to each subset of E its indicator function defines an isomorphism (of lattices) from the power set of E to the lattice $\mathcal{L}(E, \{0,1\}) = \{0,1\}^E$.

§46. Birkhoff Representation Theorem. *The functor $P \mapsto P^*$ defines a natural anti-isomorphism of categories from the category $\mathfrak{P}_{finite}$ of finite posets and order preserving maps to the category $\mathfrak{D}_{finite}$ of finite distributive lattices and lattice homomorphisms. In particular, every finite distributive lattice L is isomorphic to the lattice of down-sets of a finite poset P and the finite poset P is uniquely determined (up to order isomorphism) by L.*

Proof. See [3] pages 112-123. □

§47. Remark. Birkhoff repesentation was incorrectly called *Stone representation* in [10]. Birkhoff's version (in [1]?) dealt with the case of finite partially ordered sets whereas in [11] Stone generalized this duality to infinite partially ordered sets. See [3] and the references cited therein for an account of the history.

C First Order Logic

§48. Here's how to view an election procedure as a relational structure in the sense of model theory.[6] Let C be a finite set. Let $\mathcal{L}$ be the signature

$$\mathcal{L} := \{\Sigma, \mathsf{E}\} \cup \{\mathsf{F}_{ab}\}_{(a,b)\in C_2} \cup \{\mathsf{P}_{ab}\}_{(a,b)\in C_2}$$

where the predicates Σ, E, F_{ab} are unary and the predicates P_{ab} are binary. Let Γ_0 denote the set of universal closures of the following sentences:

$$\mathsf{\Sigma}(w) \leftrightarrow \neg\mathsf{E}(w)$$

$$\mathsf{P}_{ab}(\sigma, x) \leftrightarrow \neg\mathsf{P}_{ba}(\sigma, x)$$

$$\mathsf{P}_{ab}(\sigma, x) \wedge \mathsf{P}_{bc}(\sigma, x) \to \mathsf{P}_{ac}(\sigma, x)$$

$$\mathsf{F}_{ab}(\sigma) \to \neg\mathsf{F}_{ba}(\sigma)$$

$$\mathsf{F}_{ab}(\sigma) \wedge \mathsf{F}_{bc}(\sigma) \to \mathsf{F}_{ab}(\sigma)$$

[6] See e.g.[2] for terminology.

where (a, b) ranges over C_2. The first sentence in this list says that there are two sorts of individuals. We will use greek letters as variables ranging over one and latin letters as variables which range over the other. Thus[7]

$$\forall x\, \Phi(x) \leftrightsquigarrow \forall w \,\blacksquare\, \mathsf{E}(w) \to \Phi(w), \qquad \exists x\, \Phi(x) \leftrightsquigarrow \exists w \,\blacksquare\, \mathsf{E}(w) \wedge \Phi(w),$$

and similarly for $\forall \sigma\, \Psi(\sigma)$ and $\exists \sigma\, \Psi(\sigma)$. For $\lambda \in O(C)$ introduce the abbreviations

$$\mathsf{P}_\lambda(\sigma, x) \leftrightsquigarrow \mathsf{P}_{a_1 a_2}(\sigma, x) \wedge \mathsf{P}_{a_2 a_3}(\sigma, x) \wedge \cdots \wedge \mathsf{P}_{a_{n-1} a_n}(\sigma, x)$$
$$\mathsf{F}_\lambda(\sigma) \leftrightsquigarrow \mathsf{F}_{a_1 a_2}(\sigma) \wedge \mathsf{F}_{a_2 a_3}(\sigma) \wedge \cdots \wedge \mathsf{F}_{a_{n-1} a_n}(\sigma)$$

where $C = \{a_1, a_2, \ldots, a_n\}$ and $a_1 <_\lambda a_2 <_\lambda \cdots <_\lambda a_n$. For any two formulas $\Phi(\sigma, x)$ and $\Psi(\sigma, x)$ introduce the abbreviations

$$\Phi(\sigma) \preccurlyeq \Psi(\sigma) \leftrightsquigarrow \forall x \,\blacksquare\, \Phi(\sigma, x) \to \Psi(\sigma, x),$$
$$\Phi(\sigma) \equiv \Psi(\sigma) \leftrightsquigarrow \forall x \,\blacksquare\, \Phi(\sigma, x) \leftrightarrow \Psi(\sigma, x).$$

§49. Each election procedure $f : E \to O(C)$ determines a relational structure $\mathcal{M}_f$ of type $\mathcal{L}$ as follows. The underlying subset is the disjoint union $E \sqcup \Sigma$ and the predicate symbols are assigned values as follows:

$$\begin{aligned}
\mathcal{M}_f &\models \mathsf{E}(x) &&\iff && x \in E,\\
\mathcal{M}_f &\models \mathsf{\Sigma}(\sigma) &&\iff && \sigma \in \Sigma,\\
\mathcal{M}_f &\models \mathsf{P}_{ab}(\sigma, x) &&\iff && a <_{\sigma(x)} b,\\
\mathcal{M}_f &\models \mathsf{F}_{ab}(\sigma) &&\iff && a <_{f(\sigma)} b.
\end{aligned}$$

The following formulas abbreviate the various conditions that were imposed

[7] The notation $\leftrightsquigarrow$ means that the formula on the left is an abbreviation for the for the one on the right.

on election procedures in Section 3.

$$\mathsf{WU}_\lambda \quad\leftrightsquigarrow\quad \forall\sigma \,\blacksquare\, [\mathsf{P}_\lambda(\sigma) \equiv E] \to \mathsf{F}_\lambda(\sigma).$$

$$\mathsf{SU}_{ab} \quad\leftrightsquigarrow\quad \forall\sigma \,\blacksquare\, [\mathsf{P}_{ab}(\sigma) \equiv E] \to \mathsf{F}_{ab}(\sigma).$$

$$\mathsf{M}_{ab} \quad\leftrightsquigarrow\quad \forall\sigma\forall\tau \,\blacksquare\, \mathsf{F}_{ab}(\sigma) \wedge \textstyle\bigwedge_{c\neq b} \big[\mathsf{P}_{cb}(\sigma) \preccurlyeq \mathsf{P}_{cb}(\tau)\big]$$
$$\wedge\ \textstyle\bigwedge_{c,d\neq b} \big[\mathsf{P}_{cd}(\sigma) \equiv \mathsf{P}_{cd}(\tau)\big] \to \mathsf{F}_{ab}(\tau)$$

$$\mathsf{I}_{ab} \quad\leftrightsquigarrow\quad \forall\sigma\forall\tau \,\blacksquare\, \big[\mathsf{P}_{ab}(\sigma) \equiv \mathsf{P}_{ab}(\tau)\big] \to \big[\mathsf{F}_{ab}(\sigma) \leftrightarrow \mathsf{F}_{ab}(\tau)\big].$$

$$\mathsf{D}(z) \quad\leftrightsquigarrow\quad \forall\sigma \,\blacksquare\, \textstyle\bigwedge_{\lambda\in O(C)} \big[\mathsf{P}_\lambda(\sigma, z) \to \mathsf{F}_\lambda(\sigma)\big].$$

§50. Lemma. *A relational structure $\mathcal{M}$ of type $\mathcal{L}$ satisfies $\mathcal{M} \models \Gamma_0$ if and only if $\mathcal{M}$ is isomorphic to $\mathcal{M}_f$ for some election procedure f. Moreover for any election procedure f*

1. *f satisfies (WU) in §6 if and only if $\mathcal{M}_f \models \mathsf{WU}_\lambda$ for all $\lambda \in O(C)$.*
2. *f satisfies (SU) in §6 if and only if $\mathcal{M}_f \models \mathsf{SU}_{ab}$ for all $(a, b) \in C_2$.*
3. *f satisfies (M) in §8 if and only if $\mathcal{M}_f \models \mathsf{M}_{ab}$ for all $(a, b) \in C_2$.*
4. *f satisfies (I) in §9 if and only if $\mathcal{M}_f \models \mathsf{I}_{ab}$ for all $(a, b) \in C_2$.*
5. *f satisfies (D) in §13 if and only if $\mathcal{M}_f \models \exists z\, \mathsf{D}(z)$.*

§51. With these notations Arrow's Theorem from §14 says that every finite model $\mathcal{M}$ of the set of sentences

$$\Gamma = \Gamma_0 \cup \{\mathsf{WU}_\lambda\}_{\lambda\in O(C)} \cup \{\mathsf{I}_{ab}\}_{(a,b)\in C_2}$$

also models the sentence $\exists z\, \mathsf{D}(z)$. (Here Γ_0 is the set of sentences defined in §48)

References

[1] G. Birkhoff: On the combination of subalgebras, *Proc. Camb. Phil. Soc.* **29** (1933) 441-464.

[2] Elisabeth Bouscaren (Editor): *Model Theory and Algebraic Geometry*, Springer Lecture Notes in Mathematics **1696**, 1998.

[3] Davey, B. A and Priestley, H. A.: *Introduction to lattices and order*, Second edition. Cambridge University Press, New York, 2002.

[4] R. Dedekind: Über Zerlegungen von Zahlen durch ihre grössten gemeinsamen Teiler, Festschrift Hoch. Braunschweig u. ges. Werke(II), 1897, pp. 103-148.

[5] Jeff Kahn: Entropy, independent sets and antichains: A new approach to Dedekind's problem, Proc. Amer. Math. Soc. **130** (2002), 371-378.

[6] Alan P. Kirman and Dieter Sondermann: Arrow's theorem, many agents, and invisible dictators. J. Econom. Theory **5** (1972), no. 2, 267-277.

[7] P. Komjáth and V. Totik: Ultrafilters, Amer. Math. Monthly 115 (2008), 33-44.

[8] R. Duncan Luce and Howard Raiffa: *Games and Decisions*, John Wiley, 1958.

[9] Robert McNaughton: Unate truth functions. IRE Trans. **EC-10** (1961) 1-6.

[10] Joel W. Robbin and Dietmar A. Salamon: Lyapunov maps, simplicial complexes, and the Stone functor, Ergodic Theory and Dynamical Systems **12** (1992), no. 1, 153-183.

[11] M. Stone: Topological representation of distributive lattices and Brouwerian logics, Casopis Pest. Mat. Fys.67 (1937) 1-27.