Voting Protocols and Arrow’s Theorem

Subhash Suri

May 18, 2015

1 Voting and Social Choice Theory

Social Choice Theory is about aggregating wishes (preferences) of individuals into a societal preference. In other words, which collective (social) choice or outcome is the most logical one when the individuals of the society have conflicting preferences. Examples include choosing a president in national elections, hiring a faculty member, choosing community projects (swimming pool, ice rink, library) to fund under limited budget, etc.

2 Modeling the Problem

1. A set of \(N\) voters (agents, individuals) whose preferences are to be aggregated.

2. A set of \(n\) Candidates (outcomes, alternatives).

3. An individual’s preference is a rank ordering of all the candidates (ties allowed).

4. The goal is to find a social ordering of the candidates that is consistent with the individual preferences.

The problem is straightforward if there are only 2 candidates—majority voting works. But it becomes surprisingly more complex and subtle with 3 or more candidates. This lecture is about this complex setting.

3 Some Well-Known Voting Methods.

1. Plurality: the most preferred candidate (ranked \#1 by most voters) wins.

2. Plurality with Runoff. (Used in Presidential Elections in France.) In first round, each voter casts a single vote. Eliminate all but the top 2, then choose between those using Plurality.
3. **Single Transferable Vote (Instant Runoff).** (Used in UK.) Each agent votes for their most-preferred candidate. If there is no majority, the candidate with fewest votes is eliminated. Each voter who voted for the eliminated candidate transfers their vote to their most-preferred candidate among the remaining candidates. Repeat until a winner.

4. **Approval Voting.** (Used by Mathematical Association of America, etc.) Each voter can cast a single vote for as many candidates as he wants. The candidate with the most votes is selected.

5. **Borda Count:** Each voter rank orders all $n$ candidates. The most preferred candidate receives $n - 1$ points, the next $n - 2$ points, and the last 0. Add up these voter points for each candidate. The one with the maximum number of points wins.

6. **Pairwise Elimination:** Pair candidates with a schedule. The candidate who is preferred by a minority of voters is deleted Repeat until only one candidate is left

(Example: suppose the order is $A, B, C$. Then, first we do $A$ vs $B$. Eliminate the loser. The winner goes against $C$.)

In spite of their natural and common-sense appeal, all of these voting protocols can create counter-intuitive outcomes. For instance, the pair-wise election seems reasonable, but we can have cycles. $A$ beats $B$; $B$ beats $C$; and $C$ beats $A$.

**Condorcet Condition.** If a candidate is preferred to every other candidate in pairwise runoffs, then he/she is the winner. Important condition, but not always possible.

### 4 Counter-intuitive Behavior of Voting Protocols.

Consider the following preference lists.

$$\begin{align*}
499 & \text{ prefer } A > B > C \\
3 & \text{ prefer } B > C > A \\
498 & \text{ prefer } C > B > A
\end{align*}$$

1. Who is the Condorcet (pairwise) winner? It’s $B$ because

- $B$’s score is $1003 = (3 + 498) \text{ vs } A$ and $(499 + 3) \text{ vs } C$
- $A$’s score is $998 = 499 \text{ vs } B + 499 \text{ vs } C$
• C’s score is 999 = 3 vs A and (498 + 498) vs A and B

2. Who is the Plurality Vote winner? It’s A!

3. Who wins under Plurality with elimination (or single transferable)? It’s C!
   (B gets the least #1 votes in round 1 and is eliminated, leaving A vs C election, which
   C wins because B’s supporters switch to C.)

4. Another example to show Non-Robustness and Sensitivity

   35 prefer $A > C > B$
   33 prefer $B > A > C$
   32 prefer $C > B > A$

   (a) The Plurality winner is A. The Borda winner also is A. Calculation:

   $\begin{array}{ccc}
   & A & B & C \\
   70 & 0 & 35 \\
   33 & 66 & 0 \\
   0 & 32 & 0 \\
   \end{array}$

   (b) Now suppose C drops out. How does that affect the outcome? The new preferences

   35 prefer $A > B$
   33 prefer $B > A$
   32 prefer $B > A$

   So, now under both Borda and Plurality, B wins!

   (c) Who wins the pairwise elimination, with the order $(A, B, C)$? C wins!
      (A loses against $B$. $B$ loses against $C$.)

   (d) Who wins the pairwise elimination, with the order $A, C, B$? B wins!
      (A wins against $C$. A loses to $B$.)

   (e) Who wins the pairwise elimination, with the order $B, C, A$? A wins!!
5. Another Example:

<table>
<thead>
<tr>
<th>Preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 prefer</td>
</tr>
<tr>
<td>1 prefer</td>
</tr>
<tr>
<td>1 prefer</td>
</tr>
</tbody>
</table>

Pairwise Election with ordering: A, B, C, D.
Who wins? D!
Notice something strange. Every single voter prefers B to D, yet D wins.

5 Arrow’s Axiomatic Approach.

These paradoxical and strange outcomes of natural-sounding voting protocols should give one pause about their legitimacy, and raise the obvious question: is there is a protocol that does not exhibit these troubling behaviors? Against this backdrop, we arrive at a famous result in economics and social choice theory, called **Arrow’s Impossibility Theorem**.

Kenneth Arrow, who won the Nobel Prize in Economics, investigated an axiomatic approach to rank-order voting protocols, and set out the following 3 **Fairness Conditions**, which every useful protocol must satisfy:

1. If every voter prefer X to Y, then Social choice should also prefer X to Y.

2. If every voter’s preference between X and Y is unchanged, then the social order between X and Y should remain unchanged as well. In other words, voters’ preference changes over other candidates, U, V, Z etc, should not affect X vs Y preference.

3. There is no **dictator**: a single voter whose preferences always determines the social preference.

Remarks on Conditions:

- Condition 1 is also called **Pareto Efficiency** (PE): when all agents agree on the rank order of two candidates, then the society should respect that order. If everyone has A > B, then social order must also rank A > B.
• Condition 2 is called **Independence of Irrelevant Alternatives** (IIA): as long as all voters’ relative ranking of $X$ and $Y$ remains the same, their relative order (who comes first) should also remain the same in the social choice.

• The dictatorship conditions rules out the presence of a voter who can dictate the results, meaning that there is no single voter whose preferred order is adopted by the society **no matter** what the preferences of the remaining voters.

**Arrow’s Impossibility Theorem.** *No protocol (social choice function) can satisfy all 3 conditions: If we have PE and IIA, then the only protocol is dictatorial!*

This is both a beautiful and a highly unexpected result.

### 5.1 Proof of the Theorem.

• A **profile** is the rank ordering of the candidates by all voters (one preference list per voter).

• Intuitively, the proof uses the fact that the axioms of PE and IIA constrain how the social preference order can change as individual voter preferences change.

• We will carefully construct a set of profiles, with a very special structure, to bring out the logical inconsistency of our 3 fairness axioms.

• The proof exploits the fact that the protocol must produce outcomes that are consistent with our 3 axioms for all possible voter profiles. Starting with a highly structured profile whose outcome is indisputable, we will generate other profiles whose outcomes under the constraints of the axioms lead to contradictions.

We begin with the following Lemma that uses Pareto Efficiency.

**Pareto Lemma.** *Pick any candidate $b$. If every voter ranks $b$ either at the very top, or at the very bottom, then social order $> \text{ must put } b \text{ either at the very top or the very bottom. (No middle place.)}*


Proof of Pareto Lemma. Consider voter profiles satisfying the lemma for which $b$ is not extreme, so there are candidates both above and below in the rank order. Let $a, c$ be such that $a > b$ and $b > c$.

We now slightly modify each voter’s profile by moving $c$ just above $a$, leaving the rest unchanged. Since in each voter’s list, $b$ is at the top or the bottom, this switch does not affect the $(a, b)$ or the $(b, c)$ order. Therefore, by IIA, the social order $>$ must continue to maintain $a > b$ and $b > c$, and by transitivity $a > b > c$.

But each voter now has $c > a$, so by Pareto Efficiency, the social order must also have $c > a$. A contradiction!

Pivotal Voter Lemma. Pick any candidate $b$. There is voter $n^*$ that is pivotal for $b$. That is, while the other candidates keep their ranking of $b$, the agent $n^*$ can move $b$ from bottom to top by changing his vote.

Proof of Pivotal Voter Lemma.

1. List the voters in some arbitrary order 1, 2, ...

2. Start with a profile in which every voter puts $b$ at the bottom; remaining preferences being arbitrary.

3. By PE, the social order must also rank $b$ at the bottom.

4. Now let the voters 1, 2, ... one by one modify their profile by moving $b$ to the top, without changing any other ranking.

5. There must be one earliest voter $n^*$ whose switch causes $b$ to move to the top in the social ranking—because when all voters move $b$ to the top, social order will also have $b$ at the top, so the change must occur at some point.

6. Define $P_1$ as the profile in which the first $(n^* - 1)$ voters have $b$ at the top; rest at the bottom.

7. Define $P_2$ as the profile in which first $n^*$ voters have $b$ at the top. (This is the tipping profile.)

8. Social order for $P_1$ has $b$ at the bottom, but for $P_2$ has $b$ at the top.

9. Agent $n^*$ is the pivotal voter for $b$, and this completes the proof.

Lemma: $n^*$ is pivotal for all candidate pairs $(a, c)$ not involving $b$. 
Proof. Without loss of generality, let’s choose $a$ as candidate to consider. We construct a new profile $P_3$ by making small modifications to $P_2$, as follows.

1. First, move $a$ to the top of $n^*$’s ranking, without changing anything else. Thus, in $n^*$ ranking we will have $a > b > c$.

2. Second, let the remaining voters arbitrarily rearrange their relative rankings of $(a, c)$, but leaving $b$ in its extremal position. (That is, we want to argue no matter what their relative preference is, $n^*$ will dictate.)

3. This is our profile $P_3$.

Now let’s see how the social ordering changes.

1. In $P_1$, $b$ was at the bottom of social order, and so $a > b$.

2. In $P_3$, the relative ordering of $a, b$ are the same for all voters as in $P_1$. Therefore, by IIA, we must also have $a > b$ under $P_3$.

3. In $P_2$, $b$ was at the top of the social order, and so $b > c$.

4. In $P_3$, the relative ordering of $b, c$ are the same for all voters as in $P_2$. Therefore, by IIA, we must also have $b > c$.

Taken together and by transitivity, $P_3$ must have $a > c$ as social order. Thus, $n^*$ has the power to make an arbitrary candidate a more preferred than $c$ irrespective of the ranking of everyone else. This is the dictator, and the proof of the lemma is complete.

Unique Dictator Lemma. $n^*$ is also pivotal for any pair $(a, b)$.

Proof. The earlier discussion centered around a particular candidate $b$, but we could chosen any candidate to play the role of $b$. So, if we want to prove the pivotal role of $n^*$ on $b$, let’s consider some other arbitrary candidate $c$.

By the same reasoning as before, we must have a pivotal voter $n^{**}$ for $c$, who has the power to control the ranking for any pair $(x, y)$ not involving $c$. But $(a, b)$ is such a pair.

On the other hand, we also know that the old pivot $n^*$ is able to affect the $a, b$ ranking—for instance, it was able to change the $a > b$ ranking in profile $P_1$ to $b > a$ ranking in profile $P_2$. This is precisely how $n^*$ was identified as pivotal. So, $n^{**}$ and $n^*$ must be the same—there cannot be two dictators!
Remark: With $b$ as a fixed candidate, the pivot $n^*$ provide a vantage point of the preference space from $n^*$ perspective. It has the ability to dictate the rank order of $(a, c)$ for any pair. $n^{**}$ has a different vantage point from $c$, and it has the power to dictate the relative order of any pair such as $(a, b)$. But we know that $n^{**}$ cannot make $a > b$ under the profile $P_2$. So, it must be the case that $n^{**} = n^*$.

Thus, if social choice function obeys PE and IIA, it is dictatorial.

5.2 Interpretation and Consequences

Non-mathematical statements of the theorem such as *No voting method is fair*, or *The only voting method that isn’t flawed is a dictatorship*, are simplifications. What the theorem does say is that a deterministic preferential voting mechanism (where a preference order is the only information in a vote, and any possible set of votes gives a unique result) cannot comply with all of the conditions given above simultaneously.

Many suggest weakening the IIA criterion as a way out of the paradox. For instance, some contend IIA is an unreasonably strong criterion. It is the one breached in most useful voting systems (trivially implied by the possibility of cyclic preferences).

So, what Arrow’s theorem really shows is that any majority-wins voting system is a non-trivial game, and that game theory should be used to predict the outcome of most voting mechanisms.

This could be seen as a discouraging result, because a game need not have efficient equilibria, e.g., a ballot could result in an alternative nobody really wanted in the first place, yet everybody voted for.

For more, see Wiki.