How to Pick the Pope, or a Pizza

The Mathematics of Democracy

Brian Mintz,

Graduate Student Seminar,

Dartmouth, Fall 2024

How can we **optimally** make collective decisions?

Outline

- Background + Formalization
- Simple Majority + Condorcet’s Paradox
- Scoring Functions + Borda Count
- The Impossibility Theorems: Arrow + Sen + Gibbard-Satterthwaite…
- And how to fix them

Timeline

Greece
200 BC

Pliny
the younger,
Rome, 90

Nicolas of Cusa
1283

Jean-Charles
de Borda
1780

Marquis de
Condorcet
1785

Lewis Carroll
1876

May
1952

Arrow
1963

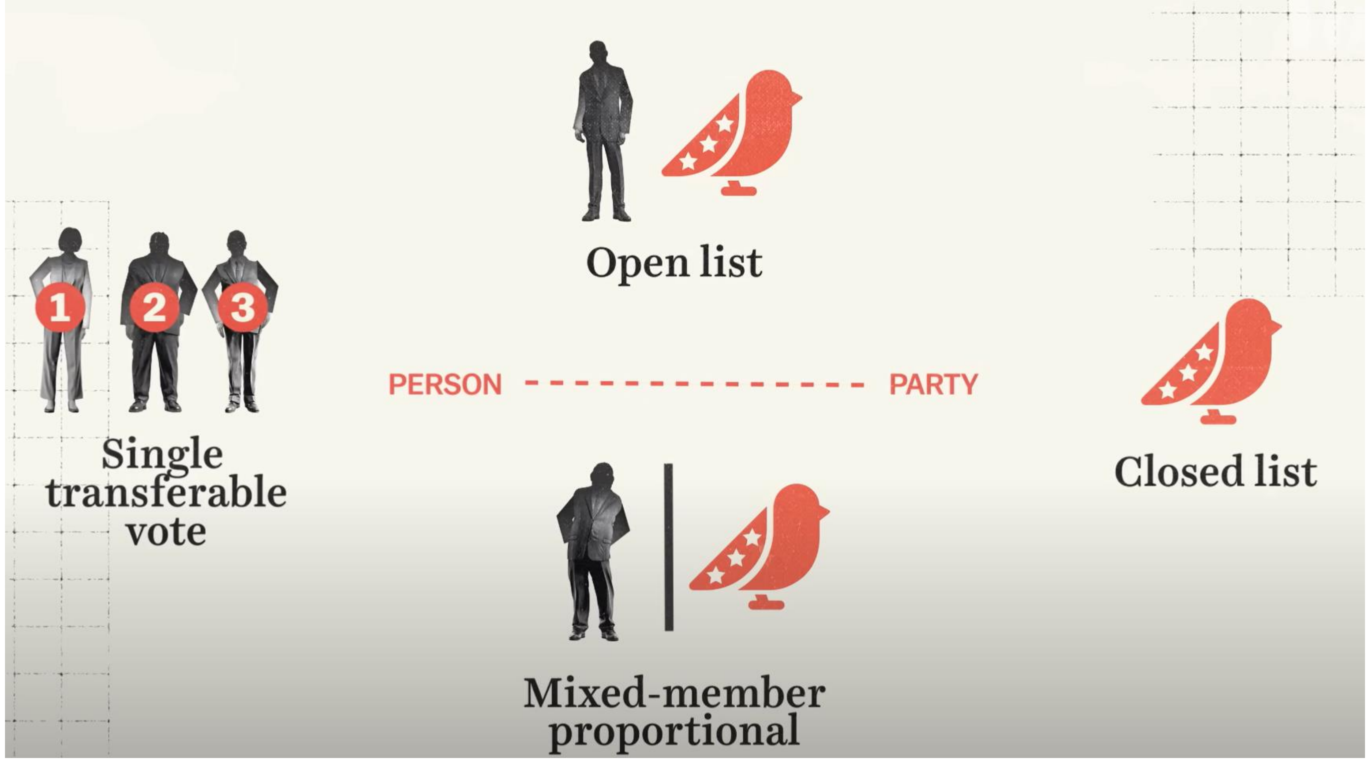
Gibbard and
Satterthwaite
1976

Us!
Today

Proportionality

Finland, Belgium, Denmark

Ireland, some local US



Spain

Germany

- Which is best, and how can we formally reason about the properties each system has?

https://www.youtube.com/watch?v=bqWwV3xk9Qk

Preferences

- A preference on a set of alternatives S is a collection of ordered pairs (x, y), written xRy, where x is "preferred" to y.
- Properties:
  - **Reflexive** – for all x we have xRx.
  - **Complete** – For any pair of elements (x, y), either xRy or yRx.
  - **Transitive** – $xRy \wedge yRz \rightarrow xRz$.
- We'll usually work with strict preferences: $xPy \leftrightarrow xRy \wedge \sim(yRx)$.
- **Theorem** [Von Neumann and Morgenstern, 1944]: **If a preference relation is reflexive and complete, then a social choice function exists if and only if the preference relation is acyclical.**

When do utility functions (numbers) suffice?

Theorem: Any complete and transitive preference relation can be represented by a utility function

Completeness allows us to compare each new alternative to all the ones already assigned a utility, to determine where to insert it.

Transitivity ensures no contradictions will occur.

$$xPy \sim x > y, \qquad xRy \sim x \geq y$$

Social aggregation and choice functions

Social aggregation function: a function from profiles of preferences to preferences (the social preference under a given profile).

Social choice function: a function from profiles of preferences to subsets of alternatives (the "winners" under a given profile).

The former gives a whole preference while the latter just gives the winner.

Simple Majority

- With this rule, the social preference is xPy if and only if the number of individuals that strictly prefer x to y is greater than the number that strictly prefer y to x.

- Example: Consider these divisions of resources among three individuals:

$$x = (2,0,1) \qquad y = (1,2,0) \qquad z = (0,1,2)$$

We have xPy, because a majority, the first and third individuals, prefer x to y.

Condorcet’s Paradox (1750’s)

Consider these divisions of resources among three individuals:

$$X = (2,0,1),\ Y = (1,2,0),\ \text{and}\ Z = (0,1,2)$$

Which alternative wins under simple majority voting?

Preferences of individuals 1, 2, and 3

1	2	3
X	Y	Z
Y	Z	X
Z	X	Y

* This preference table is a “Latin square” each alternative appears exactly once in each row and column, there’s a lot of cool math about these.

Condorcet's Paradox (1750's)

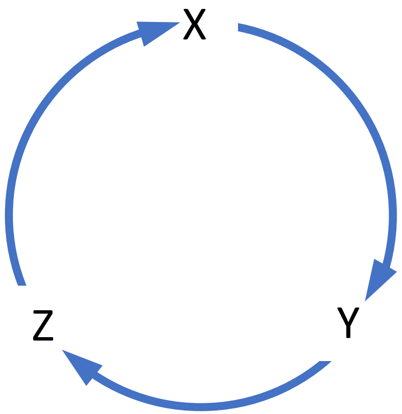
Consider these divisions of resources among three individuals:

X = (2,0,1), Y = (1,2,0), and Z = (0,1,2)

Which alternative wins under simple majority voting?

Preferences of
individuals 1, 2, and 3

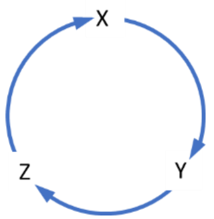
	1	2	3
	X	Y	Z
V	Y	Z	X
V	Z	X	Y



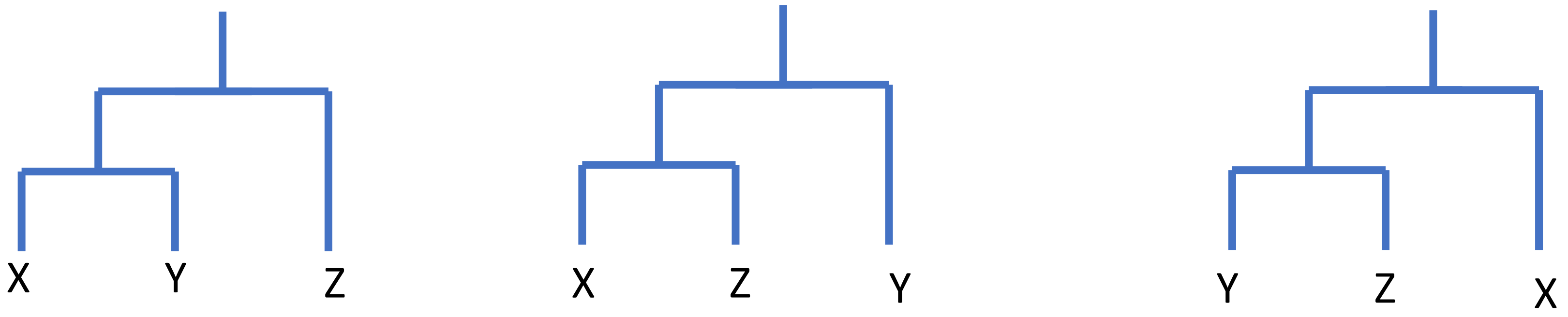
We've combined individually consistent preferences to get an nonsensical result!

* This preference table is a "Latin square" each alternative appears exactly once in each row and column, there's a lot of cool math about these.

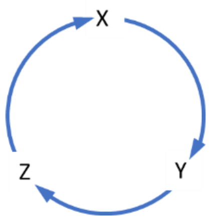
Manipulability



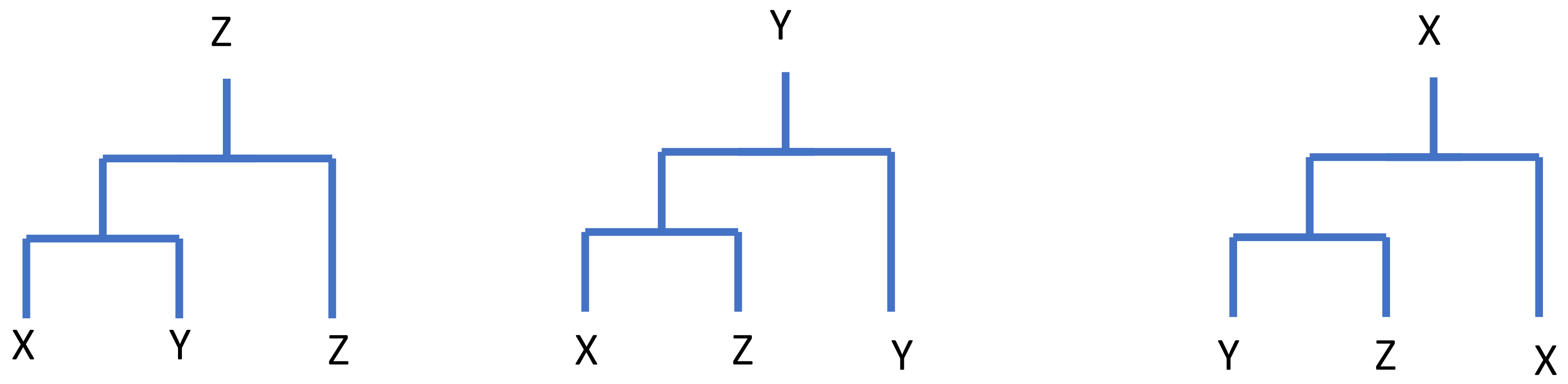
Consider the following orders / backets of votes. What is the winner in each?



Manipulability



Consider the following orders / backets of votes. What is the winner in each?



Depending on which bracket we choose, any alternative could be the winner!

The order of voting can change the outcome when the winning relation is non-transitive.

May’s Theorem (1952)

- **Anonymity** – switching about the alternatives should not change the social preference relation.
- **Neutrality** – switching about the identities of the voters should not change the social preference.
- **Positive Responsiveness** – xRy becomes xPy if one individual k changes xI_ky to xR_ky, or yP_kx to xR_ky.

May's Theorem (1952)

- **Anonymity** – switching about the alternatives should not change the social preference relation.
- **Neutrality** – switching about the identities of the voters should not change the social preference.
- **Positive Responsiveness** – xRy becomes xPy if one individual k changes xI_ky to xR_ky, or yP_kx to xR_ky.

- **Theorem: The only aggregation rule that satisfies Anonymity, Neutrality, and Positive Responsiveness is the simple majority rule.**
- Proof sketch: By A, the rule only depends on the number who are in favor or opposed. If these are equal, then the social preference must be indifference (or else there'd be a contradiction). Induction is used otherwise.

Scoring Functions

- We assume all preferences are strict linear orderings, and assign scores to each rank (first, second, third, fourth, …).
- A **scoring function** chooses the alternative(s) a_i that maximize

$$\sum_{j=1}^{n} r_j(a_i)$$

Where $r_j(a_i)$ is the score of alternative a_i for preference j.

- **Consistency**: alternatives which win two (disjoint) profiles individually should also win the union of the profiles.
- **Theorem** [Young 1975]**: The only social choice function that satisfies Anonymity, Neutrality, and Consistency are scoring functions.**

Plurality

- The score of rank one is 1, and all other ranks get score 0.

1	2	3	4
a_1	a_1	a_3	a_4
a_2	a_2	a_2	a_2
a_3	a_3	a_4	a_3
a_4	a_4	a_1	a_1

score	total
1st	a_1
2nd	a_2
3rd	a_3
4th	a_4

score	total
1	
0	
0	
0	

Plurality

- The score of rank one is 1, and all other ranks get score 0.

1	2	3	4
a_1	a_1	a_3	a_4
a_2	a_2	a_2	a_2
a_3	a_3	a_4	a_3
a_4	a_4	a_1	a_1

→

score	total
1st	a_1
2nd	a_2
3rd	a_3
4th	a_4

score	total
1	2
0	0
0	1
0	1

The social choice for this preference profile is a_1.

But what if we chose different scores for each rank?

Manipulability

- What if we used the below scores instead?

1	2	3	4
a_1	a_1	a_3	a_4
a_2	a_2	a_2	a_2
a_3	a_3	a_4	a_3
a_4	a_4	a_1	a_1

score	total
1st	a_1
2nd	a_2
3rd	a_3
4th	a_4

score	total
1	2
0	0
0	1
0	1

score	Total
3	6
2	8
1	6
0	4

score	total
1	2
0.5	2
0.25	1.75
0	1.25

score	total
1	2
0.5	2
0.4	2.2
0	1.4

- **Different scores can cause different final results!**

The Borda Count

- **Borda count:** for m alternatives, the score of rank i is $m - i$.

- **Faithful –** If the population is just one individual, then the social choice function just takes their most preferred alternative.
- **Cancellation Property –** If all pairs of alternatives are preferred equally often in each order, then the social choice is all alternatives.

- **Theorem** [Young, 1974]**: The Borda Count is the only social choice function that is Neutral, Consistent, Faithful, and has the Cancellation Property.**
- Bonus: Consistency and Cancellation guarantee Anonymity (Young).

Arrow's Impossibility Theorem (1963)

- Properties of an aggregation rule:
  - ***(U) – unrestricted domain***: any collection of preferences can be combined to get a (transitive) preference relation.
  - ***(P) – (weak) Pareto principle***: If everyone prefers x to y, then xPy.
  - ***(I) – independence of irrelevant alternatives***: The social preference between x and y depends only on the individual preferences between x and y, and no other alternatives z.

  - ***(D) – dictatorship***: The social preference is the same as that of some individual for every collection of preferences.

Arrow's Impossibility Theorem (1963)

- Properties of an aggregation rule:
  - ***(U) – unrestricted domain***: any collection of preferences can be combined to get a (transitive) preference relation.
  - ***(P) – (weak) Pareto principle***: If everyone prefers x to y, then xPy.
  - ***(I) – independence of irrelevant alternatives***: The social preference between x and y depends only on the individual preferences between x and y, and no other alternatives z.

  - ***(D) – dictatorship***: The social preference is the same as that of some individual for every collection of preferences.

- **Theorem: The only rule that satisfies (U), (P), and (I) is a dictatorship.**

Proof in a nutshell

We say a set of individuals S is "almost decisive" over a pair of alternatives (x, y), written $aD(x, y)$, if whenever everyone in S prefers x to y and everyone else prefers the opposite, then the social preference is xPy. The set is "decisive" over a pair if their preference alone determines the social preference for this pair.

1. (Contagion) If some individual is almost decisive over some pair of alternatives, then they are a dictator.

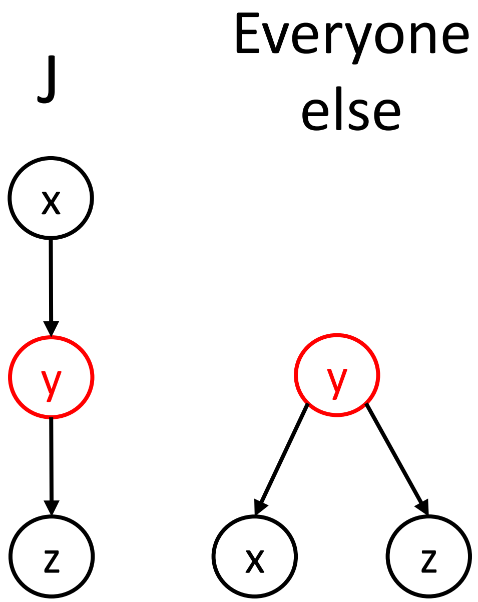
2. There is at least one individual who is almost decisive over some pair of issues.

Part 1a:

J is $aD(x,y) \to J$ is $D(x,z)$ for any z

By **(I)**, y has no effect on the social preference between x and z.

So *any preferences* between y and either x or z result in the same social preference. By **(U)**, these could be…



Since J is $aD(x,y)$, we have xPy, and by **(P)** we have yPz, so by *transitivity* xPz.

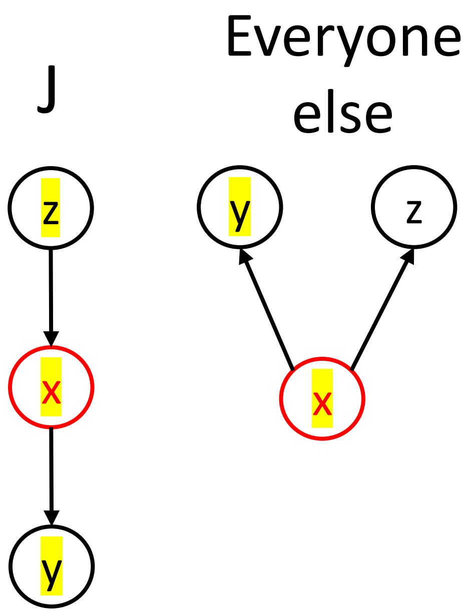
We didn't specify any preferences between x and z, *except for J's*, so J alone must be decisive for x over z.

Part 1b:

J is $aD(x, y) \rightarrow J$ is $D(z, y)$ for any z

By **(I)**, x has no effect on the social preference between y and z.

So *any preferences* between x and either y or z result in the same social preference. By **(U)**, these could be…



Since J is $aD(x, y)$, we have xPy, and by **(P)** we have zPx, so by *transitivity* zPy.

We didn't specify any preferences between y and z, *except for J's*, so J alone must be decisive for z over y.

Part 1c:

J is $aD(x, y)$
$\rightarrow J$ is $D(x, z)$ for any z (By part 1a)
$\rightarrow J$ is $aD(x, z)$ for any z
$\rightarrow J$ is $D(w, z)$ for any z and any w (By part 1b)
$\rightarrow J$ is a dictator

Intuition for part 2:

We'll consider the smallest subset V of individuals who are decisive over at least one pair of alternatives.

We'll prove by way of contradiction that V only has one member. Assuming it does not, we use **(U)** to consider the following preference profile to get a Condorcet cycle, contradicting transitivity.

Since V has only one member, they alone are almost decisive over some pair of alternatives, so are our dictator!

V_1	V_2	V_3
x	z	y
y	x	z
z	y	x

This is our (flipped) Condorcet table!

Part 2:

1. Consider all the sets that are decisive over any pair of alternatives (there is at least one, the whole population). Let V be the smallest of these sets, and (x, y) the alternatives V is almost decisive over.
2. Assume, by way of contradiction, that V has more than one member.
3. Then we can split V into V_1, some single individual, and the remainder V_2. Let V_3 be all individuals not in V.
4. Consider the preference table to the right. It will yield a preference order by **(U)**.
5. Because V is almost decisive for (x, y), we have xPy.
6. Could zPy? No, because then V_2 would be almost decisive for (z, y), contradicting our assumption that V was the smallest.
7. So yRz. **Transitivity** and step 5 give xPz. But then V_1 is almost decisive over (x, z). This is a contradiction for the same reason.
8. Since our initial assumption that V had more than one member must lead to a contradiction, it must be false. That is, V has just one member, who is therefore almost decisive over some pair of issues.
9. But by our lemma, this means that individual is a dictator!

$\overbrace{V_1 \quad V_2}^{V}$

V_1	V_2	V_3
x	z	y
y	x	z
z	y	x

This is our (flipped) Condorcet table!

Sen's impossibility of a Paretian Liberal (1970)

- **(L) - Liberalism**: For every individual i, there is at least one pair of alternatives that i is decisive over (in both directions).

Sen's impossibility of a Paretian Liberal (1970)

- **(L) - Liberalism**: For every individual i, there is at least one pair of alternatives that i is decisive over (in both directions).
- **Theorem: No social decision function can satisfy (U), (P), and (L).**

Proof: By **(L)**, let i be decisive over x and y, and j be decisive over z and w. Suppose these are all different (the other cases are handled similarly). By **(U)**, we should get a preference relation from the preferences to the right. However, we have a cycle:

$$x \rightarrow y \rightarrow z \rightarrow w \rightarrow x$$

By the decisiveness of i, property **(P)**, the decisiveness of j, and property **(P)** again (respectively). This contradiction implies no such decision function can exist!

	i	j	All other k	
	w	y	w	y
$\vee$	x	z	x	z
$\vee$	y	w		
$\vee$	z	x		

Gibbard-Satterthwaite theorem (1973)

- **Strategy-proof**: There are no sets of preferences where an individual can lie about their own preferences to cause an outcome they prefer to be selected.

- **Theorem: The only social decision function that is strategy-proof and onto,** over an unrestricted domain of strict linear preferences of at least three alternatives**, is a dictatorship.**

What do we do about these problems?

Preference cycles

- **Ignore them**: If you assume all six preference orders among X, Y, and Z are equally likely, and there are only three voters, then this paradox has ~5% chance of occurring.

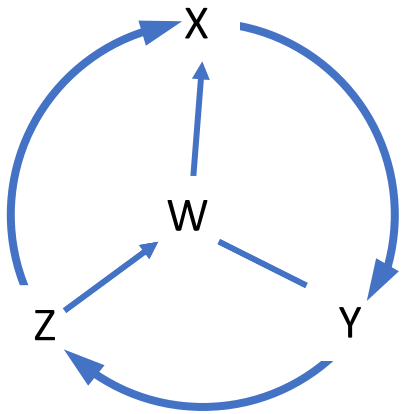
(Gehrlein, William V. "Condorcet's paradox." Theory and decision 15.2 (1983): 161-197.)

- **Accept them**: Cycles may also be fine if they are a fair depiction of the preference profile.

- **Handle them**: In such cases, one could pick one of the maximally preferred alternatives uniformly at random, or…

Copeland Method

- Choose the alternative that wins the maximum number of pairwise contests (e.g. under simple majority).
- This gives the **Condorcet winner** (also found by the Borda count).



b

a

aPb	X	Y	Z	W
X	0	1	0	0
Y	0	0	1	0
Z	1	0	0	1
W	1	0	0	0

Alternative z wins the most pairwise contests, so is the winner.

But what if there is more than one Condorcet winner?

Inversions

- Lewis Carroll [1876] *(yes the Alice in Wonderland guy!)* suggested we choose the alternative(s) that require the minimum number of preference inversions (swapping two adjacent preferences for a single individual).

- John Kemeny [1959] *(the namesake of the Dartmouth math building)* proposed a similar procedure:
  - Consider all strict linear preferences (orders) of the alternatives
  - The “total support” for an order is the sum, over all pairs of alternatives, of the number of individuals that have the same preference for that pair as the order
  - Choose the order(s) with the maximum total support.

Example

order	Pair 1 support	Pair 2 support	Pair 3 support	total
$a > b > c$	$a > b$	$b > c$	$a > c$	
$x > y > z$	$x > y$ 4+0+0	$y > z$ 0+3+0	$x > z$ 4+0+0	4+3+4=11

Preferences

1,2,3,4	5,6,7	8,9
x	y	z
z	z	y
y	x	x

Continuing in this fashion, we find the order $z > y > x$ has the most support.

This procedure agrees with the Condorcet winner / loser.

In this case, plurality would elect the Condorcet loser x.

This is the “spoiler” effect in some elections!

Full calculation

Preferences

1,2,3,4	5,6,7	8,9
x	y	z
z	z	y
y	x	x

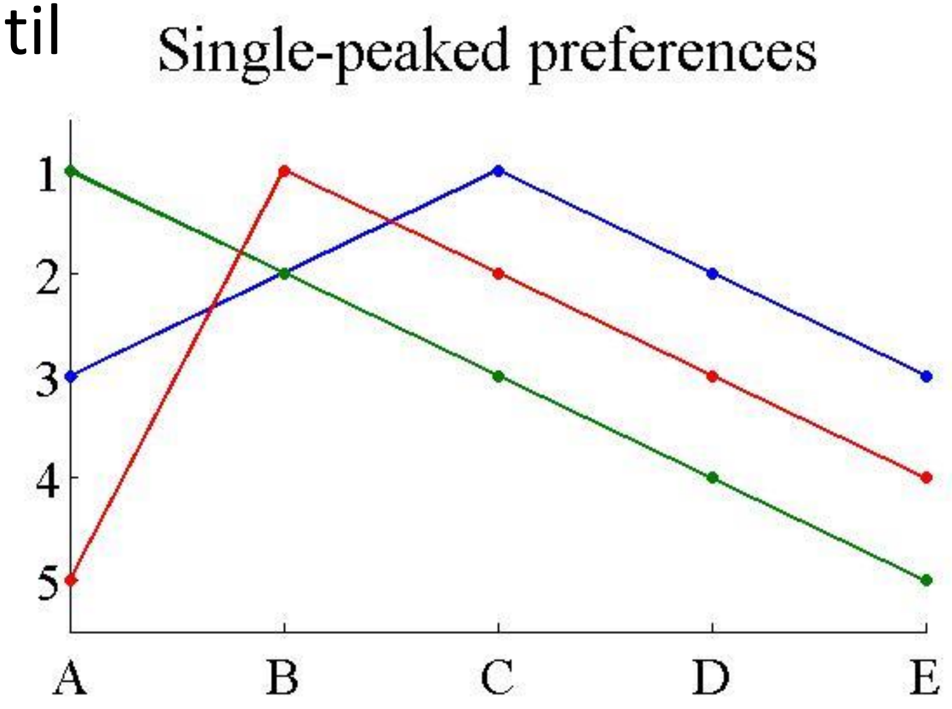
order	Pair 1 support	Pair 2 support	Pair 3 support	total
$a > b > c$	$a > b$	$b > c$	$a > c$	
$x > y > z$	$x > y$ 4+0+0	$y > z$ 0+3+0	$x > z$ 4+0+0	4+3+4=11
$x > z > y$	$x > z$ 4+0+0	$z > y$ 4+0+2	$x > y$ 4+0+0	4+6+4=14
$y > x > z$	$y > x$ 0+3+2	$x > z$ 4+0+0	$y > z$ 0+3+0	5+4+3=12
$y > z > x$	$y > z$ 0+3+0	$z > x$ 0+3+2	$y > x$ 0+3+2	3+5+5=13
$z > x > y$	$z > x$ 0+3+2	$x > y$ 4+0+0	$z > y$ 4+0+2	5+4+6=15
$z > y > x$	$z > y$ 4+0+2	$y > x$ 0+3+2	$z > x$ 0+3+2	6+5+5=16

The Impossibility Theorems

- A preference profile is **Single Peaked** if there is some ordering of alternatives where every preference increases until some peak, then decreases (with respect to this order).

- This may be realistic if there is a natural order to the alternatives,

e.g. level of military spending, size of popcorn to order.



Relaxing the (U) assumption

- **Theorem** [Black and Arrow]**: If a preference profile is single peaked,** or even just single peaked with respect to any triple of alternatives**, then simple majority voting satisfies (I) and (P) but yields a valid social preference.**

- **Median Voter Theorem** [Black, 1948]: **If the entire profile is single peaked, then the outcome of simple majority voting will be at the median, under the given order, of the set of peaks.**

Manipulability

- **Theorem** [Moulin 1980]: **If you add one less "phantom votes" than the number of voters, in any fixed positions, and assume Single Peaked Profiles, then the median voter scheme is strategy-proof (and anonymous).**
- Proof sketch: lying can only move the median further from your peak, which is less preferred, as voting anywhere on the same side of the median as your peak will not change the number of votes above and below the median.

- **Ignore it**: why is honesty even necessary, is there a problem with strategic voting? Just because there is an incentive to lie, it doesn't mean any voter has disproportionate power.

Alternative Voting Systems

- **Ranked Choice (aka Instant Runoff, aka Single Transferable Vote)**:
  - Each voter submits their ranking of the alternatives.
  - If no candidate has enough of the votes, the one with the least is eliminated and their votes are redistributed according to the rankings.
  - This repeats until one candidate has more than enough votes, and is one of the winners. Their excess votes are distributed according to the rankings
  - This process repeats until all winners are chosen (if there are more than one).

- This method has a complicated history in the U.S. (https://electionlab.mit.edu/research/instant-runoff-voting)

Pizza and The Pope

- **Approval Voting**:
  - Each voter provides the subset of candidates they approve of
  - The candidate with the highest overall approval rating is chosen.
- This avoids the spoiler effect, and was used to choose the pope between 1294-1621.

- As for pizza, you could use any of these methods. Or just order more than one topping…
- Ultimately, proportionality is difficult to achieve with single member elections. Like pizza, having multiple winners govern together is a better bet.

Summary

- By encoding preferences mathematically, we can reason about properties of different voting systems.
- A preference profile may not have a sensible aggregation.
- The voting method can affect the outcome.
- There are fundamental limits to democracy, but modifying the assumptions or more sophisticated rules can alleviate this.

What questions do you have?

References

- “A Primer in Social Choice Theory” by Wulf Gaertner
- The Mobius Strip and Chichilnisky’s impossibility theorem https://www.youtube.com/watch?v=v5ev-RAg7Xs
- Why Democracy Is Mathematically Impossible https://www.youtube.com/watch?v=qf7ws2DF-zk
- Politics in the Animal Kingdom: Single Transferable Vote https://www.youtube.com/watch?v=l8XOZJkozfI

These slides, and a lot of other cool math, are available at my website below