4. Arrow's Impossibility Theorem

Warmup

Consider the following election with 100 voters:

Number of voters

Rank	25	40	35
1	A	B	C
2	B	C	A
3	C	A	B

Q: Just considering first-place votes – what is the *societal preference order* under **plurality**?

A: The societal preference order would be
1. B
2. C
3. A

Discuss: Who does society prefer between A and B?

Q: If candidate C has to drop out the night before the election, who would win the head-to-head contest between A and B?

A: A would win 60-40.

The answer to this question ***depends on whether or not C is in the election***.

Sydney Morgenbesser and pie

There's an anecdote attributed to philosopher Sydney Morgenbesser, which captures how weird this is:

After finishing dinner, Sidney Morgenbesser decides to order dessert. The waiter tells him he has two choices: apple pie and blueberry pie. Sidney orders the apple pie. After a few minutes the waiter returns and says that they also have cherry pie at which point Morgenbesser says “In that case I'll have the blueberry pie.”

Discuss: Should your (or society’s) preference between two options be affected by the presence of a third option?

IIA

Definition

We say that a voting system satisfies ***independence of irrespective alternatives (IIA)*** if the societal preference between two candidates depends *only on those two candidates*.

As an example, we say that **plurality violates IIA**, from our example above (whether society preferred A or B depended on whether C was around).

IIA and IRV

Number of voters

Rank	25	40	35
1	A	B	C
2	B	C	A
3	C	A	B

Q: In this same election, who wins under IRV? Who wins under IRV if C drops out?

A: B wins under IRV, but A would beat B if C dropped out.

Q: Who wins under the Borda count? What about if A dropped out?

A: C wins under the Borda count. C would lose to B if A dropped out though.

So ***both IRV and Borda violate IIA***.

Suppose that one of the voters is a dictator – that is, whoever they pick wins.

Then the dictator will always pick their first choice, *regardless of who else is in the race or not.*

That is, ***dictatorships satisfy IIA.***

The Pareto Condition

Remember a voting system was ***unanimous*** if, when every voter ranks a candidate first, then that candidate wins.

We're going to slightly change this definition:

Definition

We say a voting system satisfies the ***Pareto condition*** if, when every voter ranks A over B, then in the societal preference order, A is ranked over B.

This is sometimes called ***unanimity.***

As an example, **a dictatorship satisfies the Pareto condition**. This is because the dictator's ballot *is the societal preference order*.

So if everyone ranks A over B, then the dictator ranks A over B, and therefore the societal preference order ranks A over B.

Exercise: Argue that the Borda count satisfies the Pareto condition. Show that plurality and IRV do not.

Monotonicity

Recall the definition of monotonicity:

Definition

A voting system with 2+ candidates is ***monotone*** if, for any candidate A, if some voters move A higher up their rankings, then in the resulting societal preference order, A will not decrease in ranking.

Q: Who wins this 17-person election under IRV?

A: C wins.

Number of voters

Rank	6	2	3	4	2
1	C	B	B	A	A
2	A	A	C	B	C
3	B	C	A	C	B

Q: Suppose the last column of voters swap their votes for A and C, that is, they move C even higher up their rankings. Who wins now?

A: *Shockingly, this cause B to win!*

Number of voters

Rank	6	2	3	4	2
1	C	B	B	A	C
2	A	A	C	B	A
3	B	C	A	C	B

Monotonicity

Definition

A voting system with 2+ candidates is ***monotone*** if, for any candidate A, if some voters move A higher up their rankings, then in the resulting societal preference order, A will not decrease in ranking.

In this last example, C won the election under IRV, but by moving C higher up in the rankings on some ballots and leaving everything else alone, this caused C to drop down in the societal preference order!

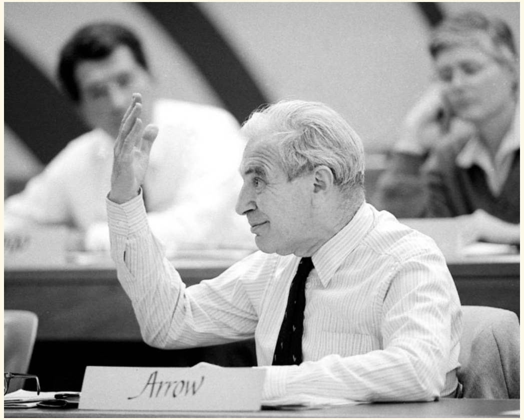
That is, ***IRV fails monotonicity***.

Voting criteria

	Plurality	Borda Count	IRV	Dictatorship
Monotone	✓	✓	🚫	✓
IIA	🚫	🚫	🚫	✓
Neutral	✓	✓	✓	✓
Pareto	🚫	✓	🚫	✓
Anonymous	✓	✓	✓	🚫

No ranked voting system (so far) has satisfied all five!

Arrow's Theorem



"No matter what I did, there was nothing that would satisfy these axioms...So after a few days of this, I began to get the idea that maybe . . . there was no voting method that would satisfy all of the conditions that I regarded as rational and reasonable. It was at this point that I set out to prove it."

- Kenneth Arrow, 1951

Arrow won the Nobel Prize in 1972 for the following theorem, which is probably the single biggest result in all of voting theory. We'll state it under slightly weaker hypotheses.

Arrow's Impossibility Theorem

If a ranked choice voting system for 3+ candidates is monotone, neutral, Pareto, and satisfies IIA, then *it must be a dictatorship*.

Arrow's Theorem

In other words, ***there is no voting system which satisfies all five conditions:***

1. Monotone
2. Neutral
3. Pareto
4. IIA
5. Anonymous

Arrow's Impossibility Theorem

If a ranked choice voting system for 3+ candidates is monotone, neutral, Pareto, and satisfies IIA, then *it must be a dictatorship*.

This tells us that in voting theory, we will never find a perfect ranked choice voting system.

We are forced to accept some imperfection, and to sacrifice some of the voting criteria we wanted to be satisfied.

Arrow's Theorem

Arrow's Theorem is true because it has a *proof*, which we will go over.

A *mathematical proof* is an airtight rigorous argument demonstrating the truth of a statement. It is how all mathematics is built.

How might we prove Arrow's Theorem?

Arrow's Impossibility Theorem

If a ranked choice voting system for 3+ candidates is monotone, neutral, Pareto, and satisfies IIA, then *it must be a dictatorship*.

We are going to *assume* we have a voting system V which is monotone, neutral, Pareto and IIA. We are then going to ***prove*** that this voting system must be a dictatorship.

The idea will be to keep changing the input ballots, and using these criteria to make arguments about how the societal preference order will change.

Proof of Arrow's Theorem

If we only assume there are 100 voters, and then we prove Arrow's Theorem in this setting, how do we still know it's true in an election with 200 voters?

In order to make sure it is *always true*, we need to make the most general statement possible. In this case we say that there are *n* voters, where *n* is any number greater than 1. If we can prove Arrow's Theorem for *n* voters, then it will be true for any value of *n*.

Suppose we have a voting system V which is *monotone, IIA, neutral, and Pareto*. By hypothesis, we have 3+ candidates, so let's name three of them – call them A, B, and C. There might be more than 3, but we won't name the other ones because we won't need to.

Proof of Arrow's Theorem

Suppose everyone ranks A first and B last. Let's write everyone's ballots from left to right. There may be a lot of them, so we'll use "…" to skip over some ballots and candidates:

Rank	Voter 1	Voter 2	…	Voter n
1	A	A	…	A
…	…	…	…	…
Last	B	B	…	B

Q: Who is ranked first in the societal preference order under V?

A: By hypothesis, V was ***Pareto***. Since everyone ranked A above any other candidate, then A will be ranked above any other candidate in the societal preference order. That is, A is ranked first.

Rank	Voter 1	Voter 2	…	Voter n	Societal Preference order
1	A	A	…	A	A
…	…	…	…	…	
Last	B	B	…	B	

Proof of Arrow's Theorem

Now suppose that Voter 1 moves B up into second place on their ranking. Does this affect the fact that A>B in the societal preference order?

A: It doesn't affect the societal preference order between A and B ***by IIA*** (since A and B are not changing in relation to each other).

Rank	Voter 1	Voter 2	…	Voter n	Societal Preference order
1	A	A	…	A	A
2	B		…		
…	…	…	…	…	
Last		B		B	

Q: Suppose Voter 1 swaps their votes for A and B. Who is ranked first in the societal preference order?

Rank	Voter 1	Voter 2	…	Voter n	Societal Preference order
1	B	A	…	A	A or B
2	A		…		
…	…	…	…	…	
Last		B		B	

A: A is preferred to *every candidate except B*, so A will either be in first or second. The only other candidate who could be in first is B.

Proof of Arrow's Theorem

If this swap put B in first place in the societal preference order, we stop there. If not, we repeat the process for Voter 2 (move B into second place, swap A and B, and see what happens).

Rank	Voter 1	Voter 2	…	Voter n	Societal Preference order
1	B	A	…	A	A
2	A	B	…		
…	…	…	…	…	
Last				B	

If we get all the way to Voter n, and A is still in first place, then once we swap Voter n's vote for A and B, we see that B *has to be in first place* by unanimity.

Rank	Voter 1	Voter 2	…	Voter n	Societal Preference order
1	B	B	…	A	A or B
2	A	A	…		
…	…	…	…	…	
Last				B	

So ***there is some voter*** (it might be Voter 1, it might be Voter 2,… it might be Voter n) for which, when we move B up to second place on their ballot and we swap A and B, the resulting societal preference order moves B into first place.

We call this person ***the pivot voter***.

Proof of Arrow's Theorem

So there is some voter, call them ***Voter k***, where 1≤k≤n, who is the ***pivot voter***. That is, when we swap A and B on their ballot, it swaps A and B in the societal preference order:

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1	B	…	B	A	A	…	A	A
2	A	…	A	B		…		B
…	…	…	…	…	…	…	…	…
3		…			B	…	B	

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1	B	…	B	B	A	…	A	B
2	A	…	A	A		…		A
…	…	…	…	…	…	…	…	…
3		…			B	…	B	

I claim that B is in 2nd place in the societal preference order before the swap.

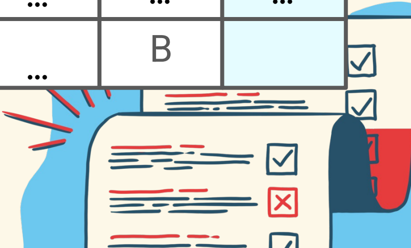
If some other candidate C is in 2nd place in the societal preference order *before the swap*, then A>C and C>B. By changing A and B on Voter k's ranking, we haven't affected A in relation to C by IIA, therefore we still have A>C, and similarly we still have C>B. But then we also have B>A, which is a contradiction.

Proof of Arrow's Theorem

So with this in mind, let's rewrite what happens under this swap on the pivot voter's ballot:

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1	B	…	B	A	A	…	A	A
2	A	…	A	B		…		B
…	…	…	…	…	…	…	…	…
3		…			B	…	B	

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1	B	…	B	B	A	…	A	B
2	A	…	A	A		…		A
…	…	…	…	…	…	…	…	…
3		…			B	…	B	



Proof of Arrow's Theorem

We remark a few things here:

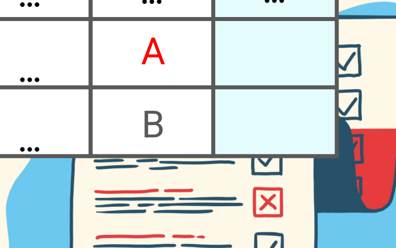
- As long as we don't move where A and B are in relation to each other, we will have that A>B still, by IIA
- As long as B doesn't move on the ballots, we will have that B is preferred to every candidate besides A.

So we are free to move A around the ballots without affecting the societal preference order by IIA!

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1	B	…	B	A	A	…	A	A
2	A	…	A	B		…		B
…	…	…	…	…	…	…	…	…
3		…			B	…	B	

So let's move A further down on everyone's ballots except Voter k:

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1	B	…	B	A		…		A
2		…		B		…		B
…	…	…	…	…	…	…	…	…
Second last		…			A	…	A	
Last	A	…	A		B	…	B	



Proof of Arrow's Theorem

From this ballot we see that A is preferred to every candidate. So as long as A doesn't move relative to any other candidate, A will remain in first place in the societal preference order.

So we can lower B in the ballots for Voters 1,2,…,k-1. This might knock B out of second place, but A will stay in first place:

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1	B	…	B	A		…		A
2		…		B		…		B
…	…	…	…	…	…	…	…	…
Second last		…			A	…	A	
Last	A	…	A		B	…	B	

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1		…		A		…		A
2		…		B		…		?
…	…	…	…	…	…	…	…	…
Second last	B	…	B		A	…	A	
Last	A	…	A		B	…	B	

Proof of Arrow's Theorem

Let's bring our third candidate, C, into the picture. Assume they are ranked 3rd last for all voters except Voter k, who has them ranked third.

Rank	Voter 1	...	Voter k−1	Voter k	Voter k+1	...	Voter n	Societal Preference order
1		...		A		...		A
2		...		B		...		?
...	...	...	...	...	...	...	...	...
Second last	B	...	B		A	...	A	
Last	A	...	A		B	...	B	

Rank	Voter 1	...	Voter k−1	Voter k	Voter k+1	...	Voter n	Societal Preference order
1		...		A		...		A
2		...		B		...		
3		...		C		...		
...	...	...	...	...	...	...	...	...
Third last	C	...	C		C	...	C	
Second last	B	...	B		A	...	A	
Last	A	...	A		B	...	B	

Proof of Arrow's Theorem

What happens if we swap B and C on Voter k's ballot?

Then C is preferred to B by every voter. So in the societal preference order, we don't know where C and B are, but we definitely know that C>B.

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1		…		A		…		A
2		…		B		…		
3		…		C		…		
…	…	…	…	…	…	…	…	
Third last	C	…	C		C	…	C	
Second last	B	…	B		A	…	A	
Last	A	…	A		B	…	B	

Rank	Voter 1	…	Voter k−1	Voter k	Voter k+1	…	Voter n	Societal Preference order
1		…		A		…		A
2		…		C		…		
3		…		B		…		
…	…	…	…	…	…	…	…	C > B
Third last	C	…	C		C	…	C	
Second last	B	…	B		A	…	A	
Last	A	…	A		B	…	B	

Proof of Arrow's Theorem

Finally, let's swap A and B in Voter k+1,...,Voter n's ballots.

Rank	Voter 1	...	Voter k−1	Voter k	Voter k+1	...	Voter n	Societal Preference order
1		...		A		...		A
2		...		C		...		
3		...		B		...		
...	...	...	...	...	...	...	...	C > B
Third last	C	...	C		C	...	C	
Second last	B	...	B		A	...	A	
Last	A	...	A		B	...	B	

Rank	Voter 1	...	Voter k−1	Voter k	Voter k+1	...	Voter n	Societal Preference order
1		...		A		...		A
2		...		C		...		
3		...		B		...		
...	...	...	...	...	...	...	...	C > B
Third last	C	...	C		C	...	C	
Second last	B	...	B		B	...	B	
Last	A	...	A		A	...	A	

A is unchanging relative to any other candidate be, so A is still preferred to every candidate ***except maybe B.***

A is still preferred to C, and C is still preferred to B. Therefore A is still preferred to B.

So we have that A is still first in the societal preference order.

Proof of Arrow's Theorem

So we have a situation where ***everyone ranks A last*** except Voter k, who ranks A first, and the societal preference order ranks A first.

We can keep naming new candidates until we run through all the candidates, and we can prove that the ***societal preference order*** is exactly the same as Voter k's ballot.

Rank	Voter 1	...	Voter k−1	Voter k	Voter k+1	...	Voter n	Societal Preference order
1		...		A		...		A
2		...		C		...		
3		...		B		...		
...	...	...	...	...	...	...	...	C > B
Third last	C	...	C		C	...	C	
Second last	B	...	B		B	...	B	
Last	A	...	A		A	...	A	

In particular, ***Voter k is a dictator***.

What we've just argued doesn't constitute a fully rigorous mathematical proof (there are a lot of details missing), but this is roughly how the argument goes.

Consequences of Arrow's Theorem

In Arrow's Theorem, we assumed that our voting system was neutral, monotone, IIA, and Pareto. It turns out that the Pareto condition is implied by the other conditions!

Lemma: If a ranked choice voting system satisfies monotonicity, IIA and neutrality, then it also satisfies the Pareto condition.

We can use this to make an even stronger version of Arrow's theorem, where we don't assume the voting system is Pareto (since it is implied by the other conditions):

Arrow's Impossibility Theorem

If a ranked choice voting system for 3+ candidates is monotone, neutral, and satisfies IIA, then *it must be a dictatorship*.

Consequences of Arrow's Theorem

Therefore Arrow's Theorem tells us that we cannot have all of the following properties in a ranked choice voting system:

1. ***Monotonicity*** (if voters change their preferences in a positive way towards a candidate, it won't cause that candidate to do worse)
2. ***Neutrality*** (elections treat candidates fairly)
3. ***IIA*** (society's preference between A and B isn't affected by candidates other than A and B)
4. ***Anonymity*** (votes are anonymous --- i.e., the voting system isn't a dictatorship).

Discuss: Which of these four conditions would you be most likely to sacrifice when choosing a voting system?

Key Vocab

- IIA (independence of irrespective alternatives)
- Pareto condition
- Arrow’s Impossibility Theorem

Key Vocab

Exercise 1: Argue that the Borda count satisfies the Pareto condition.

Exercise 2: Argue that plurality and IRV do not satisfy the Pareto condition (you have to be a little clever coming up with the right election here).

Key Vocab

Exercise 3: Prove the Lemma we stated at the end: "if a ranked choice voting system with 3+ candidates satisfies monotonicity, IIA, and neutrality, then it also satisfies the Pareto condition."

Exercise 4: Let V be a voting system with 3+ candidates, and assume V satisfies IIA and Pareto. Suppose B is some candidate, and every voter ranks B either first or last place. Then prove that the societal preference order places B either first or last (but not in between).