Math 1007: Mathematics and Politics

The George Washington University Spring 2026

Jay Daigle

Contents

1 Voting

1.1 Social Choice Functions

1.1.1 Ballots

We want to consider elections with more than 2 candidates. We call the set of candidates the *slate* and the set of voters the *electorate*. We'll generally call the candidates $A, B, C, \ldots$. (Again, instead of "candidates" these could represent bills or policies or pizza toppings, but we'll call them candidates to keep the terminology simple.)

For most of this section we'll assume each voter submits a *preference ballot*, in which they list the candidates in decreasing order of preference. So if there are four candidates $\{A, B, C, D\}$ a voter might submit a ballot that says they prefer $B > D > C > A$, which we might represent

B
D
C
A

We assume that voters are *rational*, in a specific way: we assume that each voter's preferences are *transitive*, meaning that if a voter prefers B to D and prefers D to C they must also prefer B to C. (It's not clear that people work this way in real life! But it's a convenient simplifying assumption for us to make now.)

We can represent the set of all ballots with a *profile*, similar to the profiles of section **??**. But now these profiles need to have more rows, so we might get a profile like

A	B	A	C	A	B	C
B	C	B	A	B	C	A
C	A	C	B	C	A	B

Figure 1.1

In this profile, there are seven voters. We can see the first and third prefer A to B and B to C, while the second and sixth prefer B to C and C to A. In many cases we can summarize this in a *tabulated profile*

3	2	2
A	B	C
B	C	A
C	A	B

Definition 1.1. A *social choice function* for a slate of candidates is a function that takes in a voter profile, and outputs a non-empty subset of the slate.

We think of this subset as the list of "winners" of the election. We must have at least one winner, but we can have multiple winners (which you might interpret as a tie).

There is a tremendously large set of possible social choice functions; even with just three candidates and four voters, the number of possible social choice functions is a thousand-digit number. But most of those functions aren't very interesting.

We're going to start with the most obvious social choice function, which you're probably all familiar with. But then we'll look at a few more sophisticated approaches. How to choose among them?

We'll start by looking at a simple case: the case where there are only two candidates. In section **??** we'll think through the ways we could handle an election with only two outcomes, and see that there really is an optimal answer there.

In the process, we'll identify some properties we really want a social choice function to have. In section **??** we'll look at more of those properties, especially ones that only makes sense with more than two candidates. Then in section **??** we'll try to understand how each of the methods we've looked at measure up.

1.1.2 A collection of voting methods

The Plurality Method Let's start with the obvious answer: the method used most places in the United States.

Definition 1.2. A candidate who gets more votes than any other candidate is said to have a *plurality* of the votes.

In the *plurality method* we select as the winner the candidate who is ranked first by the largest number of voters. In the case that there is a tie for the most first-choice votes, we select all the candidates who tie for the most first-choice votes.

Let's apply this to the voting profile in figure **??**. Candidate A gets 3 first-place votes, while B and C each get 2. So A is the plurality candidate and wins the election.

The plurality method is familiar and simple. It also doesn't use very much information; it only needs to know each voter's top preference (sometimes called a "vote-for-one" ballot.) In some ways, that's an advantage, since it makes voting and ballots much simpler. But it doesn't use much information about voter preferences, so it seems like we could probably produce better social choice methods somehow.

For instance, consider the following tabulated profile **??**:

In this profile, A gets 5 first-choice votes, while B, C, D each get 4 and E gets 3. So A is the plurality winner. However, we

5	4	4	4	3
A	B	C	D	E
B	C	B	B	D
C	E	D	E	B
E	D	E	C	C
D	A	A	A	A

Figure 1.2

see that all the voters who don't rank A first rank them *last*; A seems widely disliked, and probably a bad choice overall. In particular, B seems widely liked, with 4 first-place votes and *thirteen* second-place votes, out of twenty. Maybe we can develop a method that takes that other information into account?

Hare's Method/Instant Runoff Voting Let's start with a method that is currently used in a several jurisdictions. This method is currently used in Australia, and Papua New Guinea. It is also used for general elections in Alaska and Maine, and in for primary elections New York City, like the one that recently nominated Zohran Mamdani for Mayor.

Definition 1.3. *Hare's method* operates as follows. Unless all candidates have the same number of first-place votes, identify the candidate (or candidates) who have the fewest first-place votes, and eliminate them from consideration. Create a new profile where that candidate is removed, and each voter moves up their lower-ranked candidates by one place.

Repeat the process, either until only one candidate remains, or until all remaining candidates have the same number of first-place votes. The remaining candidates are the winners.

Remark 1.4. This method is frequently referred to as *Instant Runoff Voting*, *Single Transferable Vote*, or sometimes just *ranked choice voting*. The last name is common but unfortunate, since there are many voting methods that involve a ranked choice process; we'll see several in this course.

Let's apply this to the following profile:

5	4	4	4	3
B	C	A	D	E
C	A	B	A	A
E	B	E	B	B
D	E	D	E	D
A	D	C	C	C

Figure 1.3

We see that E has the fewest first-place votes, and so can be removed. That generates the profile

5	4	4	4	3
B	C	A	D	A
C	A	B	A	B
D	B	D	B	D
A	D	C	C	C

Now D and C are tied for the fewest first-place votes, with 4 each. We eliminate them both and get the profile

5	4	4	4	3
B	A	A	A	A
A	B	B	B	B

B has 5 first-place votes and A has 15, so B is removed and A is left as the winner.

But now let's go back and apply Hare's method to profile **??**. Again, E has the fewest first-place votes, and is eliminated. That gives us the following profile:

5	4	4	4	3
A	B	C	D	D
B	C	B	B	B
C	D	D	C	C
D	A	A	A	A

Now A has 5 first-place votes, D has 7, while B and C each have 4. We eliminate B and C simultaneously, giving the profile

5	4	4	4	3
A	D	D	D	D
D	A	A	A	A

So this time D wins. And this happens even though a majority prefer B to D!

Coombs's Method There are other ways to follow up on the same basic idea. Here is another way to follow the instant runoff logic that Hare's method uses:

Definition 1.5. *Coombs's method* operates as follows. Unless all candidates have the same number of last-place votes, identify the candidate (or candidates) with the most last-place votes, and eliminate them. As in Hare's method, when a candidate is eliminated, remove them from the profile and let each voter move the candidates they ranked below the eliminated candidate up a spot.

Repeat the process, eliminating candidates, until either one candidate remains, or all remaining candidates have the same number of last-place votes. The candidate or candidates who remain at the end are the winners.

We can apply this method to profiles **??** and **??**.

5	4	4	4	3
A	B	C	D	E
B	C	B	B	D
C	E	D	E	B
E	D	E	C	C
D	A	A	A	A

$\to^A$

5	4	4	4	3
B	B	C	D	E
C	C	B	B	D
E	E	D	E	B
D	D	E	C	C

$=$

9	4	4	3
B	C	D	E
C	B	B	D
E	D	E	B
D	E	C	C

$\to^D$

9	4	4	3
B	C	B	E
C	B	E	B
E	E	C	C

$\to^E$

9	4	4	3
B	C	B	B
C	B	C	C

$=$

16	4
B	C
C	B

$\to^C$

20
B

and so B wins profile **??** by Coombs's method, unlike with Hare's method.

For profile **??** we get

5	4	4	4	3
B	C	A	D	E
C	A	B	A	A
E	B	E	B	B
D	E	D	E	D
A	D	C	C	C

$\to^C$

5	4	4	4	3
B	A	A	D	E
E	B	B	A	A
D	E	E	B	B
A	D	D	E	D

$=$

5	8	4	3
B	A	D	E
E	B	A	A
D	E	B	B
A	D	E	D

$\to^D$

5	8	4	3
B	A	A	E
E	B	B	A
A	E	E	B

$=$

5	12	3
B	A	E
E	B	A
A	E	B

$\to^E$

5	12	3
B	A	A
A	B	B

$\to^B$

20
A

and so A wins.

Hare and Coombs have similar structures, but somewhat different incentives. Hare cares about the number of first-place votes, and so rewards candidates with a lot of strong support, even if they have a lot of strong detractors or haters. Coombs cares about the number of *last*-place votes, so it rewards candidates no one hates, even if no one really loves them either; it rewards middle-of-the-road moderation and widespread acceptability.

The Borda Count Method A different approach doesn't work by simulating runoff elections at all. Instead it tries to take into account all the varying levels of support a candidate has at once.

Definition 1.6. The *Borda count method* works as follows. If there are n candidates, give each candidate $n-1$ points for each voter who ranks them first; $n-2$ points for each voter who ranks them second; $n-3$ points for each candidate who ranks them third; and so on, until they get 1 point for each voter who ranks them second-to-last (and 0 points for each voter who ranks them last.)

Add up all the points; the candidate who gets the most points wins. If more than one candidate ties for the most points, all of them win.

Let's try applying this to profile **??**. There are 5 candidates, so we give 4 points for a first-place vote. We can calculate that:

- A gets $5 \cdot 4 + 4 \cdot 0 + 4 \cdot 0 + 4 \cdot 0 + 3 \cdot 0 = 20$ points;
- B gets $5 \cdot 3 + 4 \cdot 4 + 4 \cdot 3 + 4 \cdot 3 + 3 \cdot 2 = 61$ points;
- C gets $5 \cdot 2 + 4 \cdot 3 + 4 \cdot 4 + 4 \cdot 1 + 3 \cdot 1 = 45$ points;
- D gets $5 \cdot 0 + 4 \cdot 1 + 4 \cdot 2 + 4 \cdot 4 + 3 \cdot 3 = 37$ points;
- E gets $5 \cdot 1 + 4 \cdot 2 + 4 \cdot 1 + 4 \cdot 2 + 3 \cdot 4 = 37$ points.

Thus B wins by the Borda count method.

The Borda method is pretty compelling, and uses all the information from the voter profile. But it does raise some issues. For instance, if we consider figure **??**, the Borda counts are $A : 49$; $B : 54$; $C : 31$; $D : 28$; $E : 38$. So B is the Borda count winner. B is also the plurality winner, in fact. But we might notice that of the 20 voters, 15 prefer A to B; in a head-to-head matchup between the two, A would win. In fact we've seen this problem a lot; often one candidate wins despite most voters preferring a different one. We could try to build a method that takes those head-to-head matchups into account.

Copeland's Method

Definition 1.7. *Copeland's method* is the social choice function in which each candidate earns one point for every candidate they beat in a head-to-head matchup (using a simple majority method). A candidate earns half a point for every candidate they tie. The candidate with the most points at the end becomes the winner. If there's a tie for the most points, all those candidates are winners.

Again looking at profile **??**:

5	4	4	4	3
A	B	C	D	E
B	C	B	B	D
C	E	D	E	B
E	D	E	C	C
D	A	A	A	A

We need to look at each pairwise competition. For instance, we see that 5 voters prefer A to B, but 15 prefer B to A, so B wins that matchup 15 to 5, and B gets one point there. We get the following table:

	AB	AC	AD	AE	BC	BD	BE	CD	CE	D
Votes for first candidate	5	5	5	5	16	13	17	13	13	8
Votes for second candidate	15	15	15	15	4	7	3	7	7	12
Winner	B	C	D	E	B	B	B	C	C	E

We see that B gets 4 points, C gets 3 points, E gets 2 points, and D gets 1 point. So B wins by Copeland's method.

A better way to organize this is to make a *matrix*. We can make a table

	A	B	C	D	E	Total
A	-	0	0	0	0	0
B	1	-	1	1	1	4
C	1	0	-	1	1	3
D	1	0	0	-	0	1
E	1	0	0	1	-	2

In each cell, we put a 1 if the candidate whose row it is wins, a 0 if the candidate whose column it is wins, and a $1/2$ if the match is a tie. (We put a - on the diagonal, since it doesn't make sense to ask of A beats A.) Then each candidate gets the sum of their row as a score.

Now let's apply this to profile **??**:

5	4	4	4	3
B	C	A	D	E
C	A	B	A	A
E	B	E	B	B
D	E	D	E	D
A	D	C	C	C

We get

	A	B	C	D	E	Total
A	-	1	1	1	1	4
B	0	-	1	1	1	3
C	0	0	-	0	0	0
D	0	0	1	-	0	1
E	0	0	1	1	-	2

There are also some methods that are definitely *bad* ideas. These aren't serious suggestions for a democratic election; but they are by-the-definition social choice methods. These are often important as test cases—if you say something is true of "all" social choice methods, you should check in your head whether it still applies to all of these.

Definition 1.8. In the *dictatorship method*, one voter is the *dictator*. Their first-choice candidate is the unique winner.

In the *monarchy method*, one candidate is the *monarch*. That candidate is the unique winner regardless of how anyone votes.

In the *all-ties method*, every candidate is selected as a winner.

We want to spend some more time analyzing these methods of election, but there are a lot of ideas here. So we're going to jump back to an easier problem: what about elections with just two candidates?

1.2 Two-Candidate Elections

For right now we want to talk about two-candidate elections. Everything we say will also apply to votes on whether or not to pass a bill, or whether to get pizza or Chinese food; the only thing that matters is that we have exactly two choices, A and B. But for brevity we'll usually refer to them as candidates.

Obviously many elections have more than two options. (We could get Thai!) but it's convenient to start out looking at the two-candidate case. This allows us to deal with some of the issues that come up in general elections without having to think of all of them at once. This is a common technique in mathematical reasoning: whatever we want to study, we often start by working on a much simpler version where some of the difficulty disappears.

Definition 1.9. A *social choice function* for two candidates is a function whose domain is the set of all possible preferences that voters could have, and whose codomain is a set with three options: A wins; B wins; or A and B tie.

It's tempting to require social choice functions to output a definite winner, and remove the possibility of ties. But as we'll see, that can wind up being impractical. We'll return to that idea later.

1.2.1 Some methods of voting

There is a fairly obvious social choice function to start with:

Definition 1.10. The *simple majority method* for a two-candidate election is the social choice function that selects as the winner the candidate who gets more than half of all the votes cast. If each candidate gets exactly half of the votes, then the result is a tie.

So if 6 people vote for A and 4 people vote for B, then A will win. If 5 vote for A and 5 vote for B, then it will be a tie.

Weighted voting methods But note we made an implicit assumption in that last paragraph: we only gave *counts* of votes. What needs to happen for that to work? It needs to not matter who votes, just what the total is. That's true for the simple majority method, but it's not true for every possible social choice function.

Definition 1.11. A *weighted voting method* works as follows. There are n voters, and each voter is assigned a positive number of votes; we will say that voter number i has w_i votes, or a *weight* of w_i.

Set $t = w_1 + w_2 + \cdots + w_n$ be the total number of votes. A candidate who gets more than $t/2$ votes is the winner. If no candidate gets more than $t/2$ votes, the result is a tie.

Example 1.12. Suppose we have ten voters, with the following vote preferences:

A	B	A	B	A	A	A	A	B	B

We call this a *profile* of the electorate. If we don't care about who gives which vote, we can summarize this in a *tabulated profile*:

6	4
A	B

Thus we see that A gets 6 votes and B gets 4 votes. In the simple majority method, A will win.

But now suppose we have a weighted voting method where the first four voters get 5 votes each, the next three get 1 vote each, and the last three get 3 votes each. Then we have $t = 5+5+5+5+1+1+1+3+3+3 = 32$. Adding up the votes we see that A gets $5+5+1+1+1+3 = 16$ votes and B gets $5+5+3+3 = 16$ votes. That election is a tie! But note we could never have figured that out from the tabulated profile.

Poll Question 1.2.1. When does weighted voting make sense?

These weighted voting methods give a preference toward some voters over others. In the limit we can make this preference absolute:

Definition 1.13. In the *dictatorship method*, one of the voters is the *dictator*. Whoever the dictator prefers is the winner.

Example 1.14. Suppose we have a weighted voting method with ten voters, but the first voter gets twelve votes and each other voter gets one vote. Then the preference of the first voter will always win the election.

Obviously the dictatorship method is undemocratic. It is, however, mathematically a social choice function. And it does actually get used in some elections.

Example 1.15. The company Meta has two classes of stock: Class A shares get one vote per share, and Class B shares get

10 votes per share. Founder Mark Zuckerbert holds the overwhelming majority of the Class B stock, which means that while he only owns about 13% of the company, he has more than 50% of the voting power. Facebook shareholder elections are effectively run by the dictatorship method.

Supermajority Methods There's a different way we can tweak voting systems: we can have a bias toward ties.

Definition 1.16. Let p be a number such that $1/2 < p \leq 1$. The *supermajority method* with parameter p selects as the winner the candidate who gets a fraction p or more of all the votes. If there are t votes in total, a candidate must get at least $p \cdot t$ votes to win. If no candidate gets $p \cdot t$ votes, the result is a tie.

Example 1.17. A common value for p is $2/3$. If we look at the same profile above, we have ten voters and thus need $2/3 \cdot 10 = 20/3 \approx 6.66$ votes to win. A gets six votes, and B gets four votes, so the result is a tie.

The number q of votes a candidate needs to win is sometimes called the *quota* and supermajority methods are sometimes called *quota methods*. It's tempting to write that $q = pt$ but that isn't quite true, since no candidate can actually get $20/3$ votes. Instead, the quota is $q = \lceil pt \rceil$, which we read as "the ceiling of the number pt" and means the smallest integer that is greater than or equal to pt.

Poll Question 1.2.2. When would this social choice function make sense?

If $p = 1/2$, this is just the simple majority method. If $p = 1$, this is the *unanimity* method or the *consensus* method, where everyone needs to agree.

Poll Question 1.2.3. Why don't we allow $p < 1/2$?

Status quo methods A weighted voting method favors some voters; a supermajority method favors ties. But sometimes we want to favor some specific result.

Definition 1.18. Start with some social choice function (such as the simple majority method or the supermajority method), which we will call the "base method". A *status quo method* designates one of the two candidates as the status quo, and the other as the challenger. If either candidate wins under the base method, then that candidate also wins under the status quo method. If there is a tie under the base method, then the status quo method names the status quo candidate as a winner.

Example 1.19. Suppose as in example **??** we have a supermajority method with $p = 2/3$, and we designate candidate B is the status quo. We have the same voter profile, so A gets 6 votes and B gets 4 votes. The supermajority method would declare this a tie, so candidate B wins.

Poll Question 1.2.4. When does this method get used regularly? When is it a good idea?

Like with the weighted voting method, this can also be taken to an extreme:

Definition 1.20. In the *monarchy method*, one of the candidates is a *monarch*. That candidate wins regardless of how anybody votes.

Mathematically this is a *constant function*: all inputs have the same output.

Note that a monarch is different from a dictator. A dictator doesn't necessarily win the election; they just select who wins the election. A monarch doesn't get any say; they win the election whether they want to or not.

Poll Question 1.2.5. When does a monarchy method make sense?

Bloc Voting

Definition 1.21. A *block voting method* partitions the electorate into n blocs, so that every voter is in exactly one bloc. Then each bloc i is assigned a positive number w_i of votes. Each bloc conducts an election using the simple majority method, possibly with some method for resolving ties. Then the bloc casts all of its w_i votes in the main election for the candidate that won that simple majority election. The winner is the candidate who receives the most votes in the main election, with a tie if both candidates receive the same number of votes.

Example 1.22. Suppose we have the following profile, split into five blocs which each gets one vote:

A	B	A	B	A	B	A	B	B	A	A	A	B	A	B

A wins the first bloc and the fourth bloc, for 2 votes. B wins the second, third, and fifth blocs, for 3 votes. So B wins the election.

If we ignore the blocs, the tabulated voting profile here is:

8	7
A	B

Is that enough information to recover the winner of the bloc election? Who would win this election by the simple majority method?

Poll Question 1.2.6. When would this voting method make sense? Where does it get used in real life?

Testing the limits There are many, many social choice functions, and we can't possibly describe them all. But when we're analyzing social choice functions, it can be really helpful to have a list of options to consider. Especially when we're trying to prove something "always" or "never" happens, it can be useful to consider some very silly social choice functions to check our intuitions.

Two such silly methods are the dictatorship method of definition **??** and the monarchy method of definition **??**. But here are two more to consider:

Definition 1.23. In the *all-ties method*, the election is a tie, no matter how the electorate votes.

Like the monarchy method, this is a constant function.

Definition 1.24. In the *parity method*, if just one candidate gets an even number of votes, then that candidate wins. If both candidates get an odd number of votes, or both candidates get an even number of votes, then the result is a tie.

Example 1.25. Suppose we have the tabulated profile

7	4
A	B

Then B has an even number of votes and A has an odd number of votes, so B wins. If one voter who currently prefers A

switches their vote to B, we would get the tabulated profile

6	5
A	B

Now A has an even number of votes so A wins.

1.2.2 Voting method criteria

We've seen a bunch of social choice methods. We haven't said which ones are better than others, and on some level this isn't a mathematical question. We can't "prove" that weighted voting is better or worse than bloc voting, or anything like that. It depends on what you want, and what you mean by "better"!

What math can do is help you figure that out—what are your goals, and how well each method does at achieving your goals. So first we have to come up with some things we care about, and write them down as precisely as we've specified the social choice functions.

Poll Question 1.2.7. What features do you think a good voting system will have?

Followup: what kind of election were you thinking about there?

One way of thinking about our attempt to develop criteria is that there are some social choice methods that we know we don't like. For each of those, we can figure out what we don't like about it, and generate some criterion that expresses that judgment.

Definition 1.26. A method satisfies the *anonymity condition* (or *is anonymous*) if it treats all voters equally.

Another way of expressing this is that an anonymous method will always give the same result if the voters exchange ballots among themselves.

It's clear that the dictatorship method, and weighted voting methods are not anonymous. The simple majority and super-majority methods are anonymous. The monarchy, parity, and all-ties methods are also anonymous.

When we have a result we have established, we will state it as a *proposition*. Major results will be called *theorems*, and minor technical results that are useful mainly to prove other results are *lemmas*.

Proposition 1.27. *A method is anonymous if and only if its outcomes depend only on the tabulated profile.*

Proof. The phrase "if and only if" means that we have to prove two separate things. We need to show that if the method is anonymous, then it depends only on the tabulated profile; and we need to show that if a method depends only on the tabulated profile, then it is anonymous.

First let's suppose we have an anonymous method. We want to show that if two different voter profiles give the same tabulated profile, they will have the same result. Since the two profiles give the same tabulated profile, the same number of voters prefer candidates A and B respectively in each profile. That means we can start with the first profile and exchange ballots around until we have reached the second profile. Since the method is anonymous, this can't change the result; so the outcome depends only on the tabulated profile.

Conversely, suppose we have a social choice method whose outcome depends only on the tabulated profile. If we have two profiles that differ only by exchanging ballots between voters, they will produce the same tabulated profile and thus produce the same result. Therefore, the method must be anonymous.

□

Definition 1.28. A method satisfies the *neutrality criterion* or *is neutral* if it treats both candidates equally.

Status quo methods and monarchy are not neutral. Majority and supermajority methods are neutral, as are weighted voting methods and dictatorships. The parity and all-ties methods are neutral.

So far simple majority and supermajority satisfy both our criteria, as do parity and all-ties. But parity, in particular, is obviously absurd. We should come up with a criterion that explains why parity is a bad voting method. And there is in fact a criterion, which feels so obvious that you might not think to state it if you're not confronted with something as silly as parity. (But we'll see that it can be trickier than it sounds in multi-candidate elections!)

Definition 1.29. A method satisfies the *monotonicity criterion* or *is monotone* if a candidate is never hurt by getting more votes.

That is: suppose the votes are cast and the method selects one candidate as the winner. Suppose the method is applied again after one or more voters change their votes from the losing candidate to the winning candidate. The candidate who was the winner before the change must remain the winner after the change.

Monotonicity is necessary if we want to avoid strategic voting: in a monotone social choice method, it's always reasonable to vote for the candidate you actually prefer. In a two-candidate election any reasonable method will be monotone, but not every method will be.

Proposition 1.30. *The parity method violates monotonicity.*

Proof. This proposition only requires us to present a *counterex-*

ample. To say a method *is* monotone is to say that a certain thing always happens. To show it is not monotone, we don't need to show that the thing never happens; we only need to show that it doesn't always happen. (Note that "it doesn't always rain" is very different from "it never rains"!)

Consider the profile

A	B	A	B	A	A	A	A	B

In this profile, A gets 6 votes and B gets 3 votes, so A wins.

Now suppose the second voter changes their mind and starts to prefer A. We get the following profile:

A	**A**	A	B	A	A	A	A	B

Now A gets 7 votes and B gets 2. By the parity method, B wins this election.

Thus a voter changing their vote from B to A causes B to win, and the parity method violates monotonicity.

□

The all-ties method also seems bad. What's wrong with it? Well, it always produces ties. A voting method should, ideally, give us a result.

Definition 1.31. A method satisfies the *decisiveness criterion* or is *decisive* if it always chooses a winner, that is, never produces a tie.

Poll Question 1.2.8. Which of the methods we've discussed so far are decisive?

This criterion is ideal, but it's actually too strong to be useful. Sometimes a tie is the only really reasonable output, like if both candidates get exactly the same number of votes. (We'll make this claim precise in section **??**). So we want a slightly weaker criterion that isn't too strong but still rules out nonsense like the all-ties method.

Definition 1.32. A method satisfies the *near decisiveness criterion* or is *nearly decisive* if ties can only occur when both candidates receive the same number of votes.

These two criteria aren't exclusive. Any decisive method is also nearly decisive. Note that this isn't normally how we use English; but it is definitely the way that we wrote these definitions. However, many methods are nearly decisive, but not decisive.

1.2.3 May's Theorem

Proposition 1.33. *The simple majority method is nearly decisive.*

Proof. Suppose the number t of voters is odd. No candidate can receive exactly half the votes, since $t/2$ is not an integer.

So one must receive more than half; they win a majority and win the election.

Now suppose t is even. If both candidates get $t/2$ votes, they receive the same number of votes, and tie. If not, one gets more than $t/2$ and thus has a majority and wins. So ties only occur when each candidate gets exactly $t/2$ votes. □

So far we haven't found any method that checks all our boxes. The simple majority method checks most of them, but it isn't decisive. However, Kenneth May in 1952 proved that if we want to check off all our criteria, we're asking for too much.

Theorem 1.34 (May's Theorem)**.** *In an election with two candidates, the only voting method that is anonymous, neutral, monotone, and nearly decisive is the simple majority method.*

Proof. Suppose we have an anonymous, neutral, monotone, nearly decisive social choice function for two candidates. Since it is anonymous, we only need to consider tabulated profiles by proposition **??**. Then let a be the number of voters who support candidate A, and b the number of voters who support candidate B; set $t = a + b$ to be the total number of voters. We want to show that the method we are imagining must be the simple majority method.

First consider the case where t is even. If $a = b = t/2$ then each candidate gets the same number of votes, so by neutrality the result must be a tie, which is the result that the simple majority method would give.

Suppose candidate A has a majority, that is, $a > t/2$. We want to show that for *any* method that is neutral, monotone, and nearly decisive, A must win. By near decisiveness, the result cannot be a tie, so either A wins or B wins.

But B can't win this election. We see that B gets $b = t - a$ votes, and since $a > t/2$ then $b < t/2$. If B wins with $b < t/2$ votes, then by monotonicity, B must win with $t/2$ votes. But we saw that if B gets $t/2$ votes then the election is a tie.

So if A gets a majority, we've seen the election can't be a tie and B can't win; thus A wins. By neutrality, the same is true of B; so in either case, a candidate with a simple majority of votes must win.

Now consider the case where t is odd. In this case, neither candidate can get $t/2$ votes, since it's not an integer; by near decisiveness, some candidate must win.

Suppose A gets $a > t/2$ votes. By near decisiveness, this can't be a tie. We claim that B cannot be the winner. We see that B gets $b = t - a < t/2$ votes, and necessarily $b < a$. If B wins with $b < a$ votes, then by monotonicity B would also have to win with a votes. But by neutrality, that means that A would also win with a votes.

So if A gets $a > t/2$ votes, then the election isn't a tie and B doesn't win; so A must win the election. By neutrality, if B gets more than $t/2$ votes, then B must win as well. So a candidate with a simple majority of votes must always win. □

The important thing about this theorem is it doesn't just prove that we haven't come up with a better method than simple majority. We've shown that any method that satisfies these criteria has to be the simple majority method; there's no possible way to be "more clever" and come up with a better option. In particular, we have the following *impossibility result*:

Corollary 1.35. *It is impossible for a voting system with two candidates to be anonymous, neutral, monotone, and decisive.*

Proof. If a method is decisive, it must be nearly decisive. If a method is anonymous, neutral, monotone, and nearly decisive, it must be the simple majority method by Theorem **??**. So any method fitting these criteria must be the simple majority method; but the simple majority method is not decisive, so no method fits all these criteria. □

A similar argument to the proof of May's theorem can give us the somewhat more general result below:

Theorem 1.36. *In an election with two candidates, a voting method that is anonymous, neutral, and monotone must be the simple majority method, a supermajority method, or the all-ties method.*

We're not going to prove this right now but I encourage you to think about why that should be true, and how you'd prove it.

1.3 Voting System Criteria

We've now fully resolved the topic of two-candidate elections. We've seen that the Simple Majority method is, in a real sense, the best method to handle such elections. (Again, depending on what you *want*; there are times when a supermajority method, or a weighted voting method, are extremely reasonable answers to specific questions.)

Now we want to take the ideas and skills we developed working on the two-candidate case, and apply them to multi-candidate races. That means we need to lay out some criteria we'd like our multi-candidate voting systems to satisfy.

We can start with an easy and obvious criterion:

Definition 1.37. *Unanimity criterion* or *unanimous* if whenever all voters place the same candidate at the top of their preference orders, that candidate is the unique winner.

This is obviously desirable, but so obvious that it's maybe not useful. (If we expected all voters to have the same first choice, we wouldn't need to worry too much about voting systems.)

In the other direction, here's something that's desirable but much too hard to get:

Definition 1.38. A method is *decisive* if it always selects a unique winner.

As we saw in section **??**, it's very hard to get a method that's decisive, and probably not a good idea. Sometimes a tie really is the "correct" answer.

Definition 1.39. A method satisfies the *majority criterion* if, whenever a candidate receives a majority of the first-place votes, that candidate must be the unique winner.

Definition 1.40. A method is *anonymous* if the outcome is unchanged whenever two voters exchange their ballots.

Lemma 1.41. *A social choice function is anonymous if and only if it depends only on the tabulated profile.*

Proof. See proposition **??**. □

Definition 1.42. A method is *neutral* if it treats all candidates the same, in the following sense: suppose we have some profile that names A to be a winner. Now suppose there is another candidate B, and all voters exactly swap their preferences for A and B. In the new profile, B should be a winner.

Also should be the case that voters liking you more should never hurt you.

Definition 1.43. A method is *monotone* if: suppose there is a profile in which candidate A wins, but some voter puts another candidate B immediately ahead of A. If that voter moves A up one place to be ahead of B, then the method must declare A to be a winner in the new profile as well.

It follows that if A moves up any number of places for any number of voters, they still must win. This is extremely desirable because it rules out "tactical" voting: if you like A more than B you should always rank A over B. Ranking A over B can never cause A to lose.

Discussion Question 1.44. Consider the following voter profile:

A	C	B
B	A	C
C	D	D
D	B	A

Who *shouldn't* win?

Definition 1.45. A method is *Pareto* or satisfies the *Pareto criterion* if whenever every voter prefers a candidate A to another candidate B, then the method does not select B as a winner.

Named after Italian economist Vilfredo Pareto. A choice is "Pareto optimal" if there's no other choice that's better for everyone. This guarantees that a winner is Pareto optimal. (Note there can be many Pareto optimal outcomes.)

Does not mean that if everyone prefers A to B then A must be a winner! Maybe everyone prefers C to both, as in this example:

7	6
D	C
C	D
A	A
B	B

Definition 1.46. A candidate is a *Condorcet candidate* if they beat every other candidate in a head-to-head. They are an *anti-Condorcet candidate* if they lose to every other candidate in a head-to-head.

Not always a Condorcet candidate. Famously

A	B	C
B	C	A
C	A	B

A beats B, B beats C, but C beats A.

People sometimes call Copeland's method the Condorcet method, but that's not quite right. And people sometimes say that the Condorcet method picks the Condorcet winner, but that's not actually a method—doesn't always give an answer!

Definition 1.47. A method satisfies the *Condorcet criterion* if whenever there's a Condorcet candidate, they're the unique winner.

A method satisfies the *anti-Condorcet criterion* if whenever there's an anti-Condorcet candidate, they don't win.

Definition 1.48. A method is *independent* if:

Suppose there are two profiles where no voter changes their mind about whether candidate A is preferred to candidate B: if a voter ranks A above B in the first profile, they must also rank A above B in the second profile. Suppose that in the first profile, the method makes A a winner but not B. Then the method must not choose B as a winner for the second profile.

The idea here is that to decide if A beats B, it shouldn't matter what voters think about another candidate C. So if we consider the two profiles:

A	C	B	C	C	A	B
B	A	C	B	B	B	A
C	B	A	A	A	C	C

A	C	B	C	C	A	B
B	A	C	B	B	B	C
C	B	A	A	A	C	A

The only difference between the two is that in profile 1, the last voter prefers A to C, while in profile 2 they prefer C to A; preferences between A and B are unchanged. If A wins and B loses in the first profile, that shouldn't change in the second!

Unfortunately, that's much harder than it sounds.

Poll Question 1.3.1. What methods have we looked at violate this?

Proposition 1.49. *Any social choice function that satisfies anonymity and neutrality must violate decisiveness.*

Proof. Since it's anonymous, we only have to look at tabulated profiles. Suppose we have $2n$ voters.

n	n
A	B
B	A

n	n
B	A
A	B

By neutrality, if A wins in the first B must win in the second, and vice versa. So both have to win.

□

Proposition 1.50 (Taylor)**.** *No social choice function involving at least three candidates satisfies both the independence criterion and the Condorcet criterion.*

Proof. Suppose we have an independent Condorcet method. Suppose there are three candidates and three voters, and we have the profile

A	C	B
B	A	C
C	B	A

We see that A can't be a winner. Because if we look at

A	C	C
B	A	B
C	B	A

C is the Condorcet candidate, and so must be the unique winner, and so A is not a winner. But since this doesn't change the relative positions of A and C, by independence A can't win in the first profile either.

We can make the exact same argument to show that B and C also can't be named winners in the original profile. But every profile must have a winner. So no method can be both Condorcet and independent.

□

Proposition 1.51. *If a method is Condorcet then it satisfies the majority criterion.*

Proof. Suppose A has a majority of first-place votes. Then they will win any head-to-head matchup, and thus A is the Condorcet candidate. So any method that satisfies the Condorcet criterion will cause A to win, also satisfying the majority criterion. □

In this sense we can say the Condorcet criterion is stronger than the majority criterion.

1.4 Evaluating Social Choice Functions

To summarize, we have discussed the following social choice functions:

- Plurality
- dictatorship
- monarchy
- a
- Borda count
- Hare's method
- Coombs's Method
- C
- Sequential agenda
- Positional methods

We also have several criteria we looked at:

- unanimous
- decisive
- majoritarian
- anonymous
- neutral
- monotone
- Pareto
- condorcet
- anti-condorcet
- independent

We want to think about how each of these criteria apply to each of the social choice functions we've studied.

1.4.1 Plurality method

Proposition 1.52. *The plurality method is majoritarian, monotone, and Pareto, but not Condorcet, anti-Condorcet, or independent.*

Proof. The majority is always a plurality, so the plurality method is majoritarian.

It's monotone because raising a candidate on some preference lists can't reduce their number of first-place votes, or increase any other candidate's. So if a candidate wins before the raising they will still win after.

It's Pareto because if A is ahead of B on every preference list, then B will have no first-place votes at all and thus will not win.

But consider the profile in figure **??**:

2	3	2
A	B	C
C	A	A
B	C	B

Figure 1.4

In this profile, B wins the plurality vote. But A beats B in a head-to-head (four to three), A beats C head-to-head (five to four), and C beats B head to head (four to three). So A is the Condorcet candidate, and does not win; B is the anti-Condorcet candidate, and yet wins. So plurality fails both Condorcet and anti-Condorcet.

Now consider the following two profiles, "before" and "after".

A	A	A	A	B	B	B
B	B	C	C	A	A	A
C	C	B	B	C	C	C

$\rightarrow$

A	A	C	C	B	B	B
B	B	A	A	A	A	A
C	C	B	B	C	C	C

In the first profile, A wins, with four first-place votes. In the second profile, two voters have changed their minds and now rank C above A; this causes B to win.

(This is exactly the sort of "strategic voting" we see in a lot of primary elections, but also in some general elections. Imagine that A is the Democrat, B is the Republican, and C is a far-left candidate. People will say that voting for C is "throwing away your vote" and just giving the election to B.)

□

1.4.2 Antiplurality method

Another method we could consider is the *antiplurality method*.

Definition 1.53. The *antiplurality method* names as winner the candidate with the fewest last-place votes.

Proposition 1.54. *The antiplurality method is monotone, but not majoritarian, Condorcet, anti-Condorcet, Pareto, or independent.*

Proof. It's monotone because raising a candidate in a preference list can never increase their number of last-place votes, or decrease the last-place votes of any other candidate. If A wins before the change, they will also win after.

Consider the profile In this profile, B has a majority first-

C	C	B	B	B
A	A	C	A	A
B	B	A	C	C

Figure 1.5

place votes, but also a plurality of last-place votes, and so loses under anti-plurality and so the method is not majoritarian.

Similarly, we see that B is the Condorcet candidate, and A is the anti-Condorcet candidate. But A wins and B loses. That means this method is neither Condorcet nor anti-Condorcet.

We can also construct a profile where A is rated higher than B by all voters, but B still wins. That's a little bit weird, because it means A can't have any last-place votes; but B can tie with

A as a winner by not having any last-place votes either. So this method isn't Pareto.

A	A	A
B	B	B
C	C	C

Finally, we want to think about independence. Consider the following two profiles:

C	A	A	B	B
A	B	B	A	A
B	C	C	C	C

$\rightarrow$

C	A	A	B	B
A	B	B	C	C
B	C	C	A	A

Figure 1.6

In the first profile, A wins the anti-plurality election, and B loses (along with C). In the second profile, we have swapped some voters between C and A, but keep the same relative positions of A and B; now B wins the anti-plurality election, and A And C both lose.

□

1.4.3 Hare's method

One of the most popular methods is Hare's method. But it is actually surprisingly weak by our criteria, and in particular it's the first method we're going to talk about that, surprisingly, isn't monotone.

Proposition 1.55. *Hare's method is majoritarian and Pareto, but not monotone, Condorcet, anti-Condorcet, or independent.*

Proof. A candidate at the top of a majority of preference lists will always win in Hare's method, because they will never have less first-place votes than anyone else and so will never be eliminated.

Conversely, if a character is not at the top of any preference lists, they will always be eliminated in the first round. If everyone prefers A to B then B will not appear at the top of any preference lists and so will be immediately resolved.

The following pair of before-and-after preference lists shows that Hare's method violates monotonicity:

6	5	4	2
A	C	B	B
B	A	C	A
C	B	A	C

$\rightarrow$

6	5	4	2
A	C	B	A
B	A	C	B
C	B	A	C

Figure 1.7

In the "before" profile, Hare's method will first eliminate C, and then in a head-to-head A will defeat B with eleven votes to six. In the "after" profile, two voters have switched their votes to prefer A to B. this means that in the first round, B is eliminated, and then in the second round C defeats A nine to eight.

To see that Hare violates Condorcet and anti-Condorcet, we can look back to figure **??**:

2	3	2
A	B	C
C	A	A
B	C	B

In the first round, A and C are eliminated, leaving B as the unique winner. But A is the Condorcet winner and B is the Condorcet loser.

Finally, to consider independence, consider the following tabulated profiles:

2	2	1
B	A	A
A	C	B
C	B	C

$\rightarrow$

2	2	1
B	C	A
A	A	B
C	B	C

In the "before" figure, C is eliminated in the first round, then B is eliminated in the second round, leaving A as the unique winner. In the "after" profile, two voters have switched their preferences on A and C, but this leaves the relative rankings of A and B unchanged. But now, A loses in the first round, and in the second round C is eliminated leaving B as the unique winner. This violates independence.

□

1.4.4 Coombs's method

Proposition 1.56. *Coombs's method is Pareto but not majoritarian, monotone, Condorcet, anti-Condorcet, or independence.*

Proof. If candidate A is ahead of candidate B on every preference list, then A will never have any last-place votes until B is eliminated. Therefore A will survive each round, with no last place votes, until B is eliminated; the election cannot come to an end until that happens. So B must lose, and Coombs's method is Pareto.

We saw in figure **??** that a candidate can have a majority and still lose an antiplurality election:

C	C	B	B	B
A	A	C	A	A
B	B	A	C	C

Here B is the majority winner, but is eliminated in the first round of Coombs's method and thus does not win. This method violates the majority criterion, and thus also the Condorcet criterion.

It *seems* like Coombs's method should satisfy the anti-Condorcet criterion, since even if they survive to the final round, they will lose the final head-to-head matchup (since they lose any head-to-head matchup). But this is not true! If we look at this same profile **??** again, we can see that A is the anti-Condorcet candidate losing to both B and C. But A wins the election, precisely because there is no "final" head-to-head: B and C are both eliminated at the same time, leaving A the unique winner.

The remaining two properties we leave as an exercise for the student. □

Exercise 1.57. *You should prove that Coombs's method isn't monotone or independent. Each of these arguments requires generating an example pair of profiles. You may wish to use the proof of proposition* **??** *as a hint or guideline as you construct these.*

1.4.5 Borda count

Proposition 1.58. *The Borda count method is monotone, anti-Condorcet, and Pareto, but not majoritarian, Condorcet, or independent.*

Proof. Monotone because raising A on some preference lists will never decrease A's score, or increase anyone else's. So if A wins before the change, they will also win after.

It's Pareto because if A is ranked above B by every voter, then A will definitely get a higher score than B, and so B will not win.

The anti-Condorcet argument is kind of tricky. (Note this is the first time we're showing something *is* anti-Condorcet, so this will be a new type of argument.) Suppose that A is the anti-Condorcet candidate; we need to show that it's not possible for them to win. That means it's not enough to just give an example; we need to give a totally universal argument.

Instead, we need to make a tricky argument with counting. Let's suppose there are n candidates and m voters. Then every voter gives out $\frac{n(n-1)}{2}$ points, and the total number of points is $\frac{mn(n-1)}{2}$. That means the *average* number of points each candidate can get is $\frac{m(n-1)}{2}$ (the total divided by the number of candidates).

The most total points any candidate can get is $m(n-1)$, since there are m voters and a first-place vote gives $n-1$ points. We want to show that the anti-Condorcet candidate A gets less than half the maximum number of points.

Another way of thinking about the Borda count is that A gets one point for each pair (X,v) of (candidate, voter) such that the voter v ranks candidate X below A. That is, A gets one point for each time a voter ranks a candidate below them. But since A is the anti-Condorcet candidate, they are ranked below each candidate more often than they are ranked above that candidate—that's what it means to lose each round head-to-head. So A gets less than half the maximum number of possible points: less than $\frac{m(n-1)}{2}$. This is precisely the average number of points a candidate gets.

Since A gets less than the average, someone must get more than average. Thus someone gets more points than A, and A cannot win.

Now we want to show that the Borda count isn't majoritarian, Condorcet, or independent. For these we can simply show examples again.

Consider the profile

A	A	A	B	B
B	B	B	C	C
C	C	C	A	A

In this case, A is the Condorcet candidate, with a majority of the first-place votes. But A only gets 6 Borda points, while B gets 7, so B wins the Borda count election.

The profiles in figure **??** also show that the Borda count isn't independent. In the first profile A gets 7 points, B gets 6, and C gets 2, so A wins. In the second profile, we've only changed the relative positions of A and C; but now A gets 5 points, B still gets 6, and C gets 4. Now B wins.

□

1.4.6 Copeland's method

The two elimination methods were very similar; the positional methods are very similar. Copeland's method is structurally quite different and therefore produces a very different set of results. In particular, this system is constructed almost specifically to satisfy the Condorcet property.

Proposition 1.59. *Copeland's method is majoritarian, Condorcet, anti-Condorcet, monotone, and Pareto, but not independent.*

Remark 1.60. We'll see later than while you can different tradeoffs than this, you can't do *better* than this, in a specific way.

Proof. In the Copeland system, a Condorcet candidate gets a perfect, maximum score, since they beat each other candidate head-to-head. Each other candidate loses at least one head-to-head and so does not get a perfect score. So the Condorcet candidate always wins. This also means the system satisfies the majority property, since a majority candidate will be a Condorcet candidate.

An anti-Condorcet candidate loses each head-to-head and thus gets a score of zero. Since some candidate will have a score greater than zero, the anti-Condorcet candidate will always lose.

We now need to prove this method is monotone and Pareto. Again, this will require a bit of a sophisticated argument; we can't just generate an example, but need to provide an argument that will work for any possible profile, no matter how strange.

To prove Pareto: suppose A is above B on every preference list. In this case, B may win some head-to-head matchups, but any matchup B wins, A will also win. This means that A gets a point every time B gets a point, and A gets at least half a point every time B gets half a point. Further, A will defeat B, so A will have at least one more point than B does. So B cannot win.

To prove monotonicity: moving candidate A up on a preference list can never hurt A in any head-to-head matchup, so it cannot reduce A's score. A change that only involves moving A up can't affect any other head-to-head matchup, so it won't ever increase any other candidate's score. Therefore, if A was a winner before the change, they will still win after the change.

But even this argument may suggest why Copeland's method isn't independent. An improvement in A's place can't improve the scores of other candidates; but changing the relative positions of other candidates can change *their* scores. So consider the following two profiles:

B	A	A		B	A	C
C	B	C	$\to$	C	B	A
A	C	B		A	C	B

Figure 1.8

In the first profile, A gets 2 points, and B gets one, so A is the unique winner. In the second profile, we have swapped the positions of A and C for one voter. Now C beats A, getting one point; A beats B, getting one point; and B beats C, getting one point. All three candidates are declared winners.

□

1.4.7 Black's method

Here is a new method that also tries to prioritize Condorcet winners.

Definition 1.61. *Black's method* is the social choice function that chooses the Condorcet candidate as the unique winner if there is a Condorcet candidate, and chooses the Borda count winner if there is not.

We can also think of this as taking the Borda count, and "fixing" the "glitch" where it doesn't satisfy the Condorcet property.

Proposition 1.62. *Black's method is majoritarian, Condorcet, anti-Condorcet, monotone, and Pareto, but not independent.*

Proof. By definition, if there's a Condorcet candidate, Black's method selects them as the unique winner. So Black's method is Condorcet, and therefore majoritarian.

We know that an anti-Condorcet candidate can't be the Condorcet candidate, and we've also see that they can't be a Borda count winner in the proof of proposition **??**. So an anti-Condorcet candidate can't win in Black's method.

To see Black's method is monotone, we need to split into two cases. Suppose the Black winner was a Condorcet winner. Then moving them up in some rankings will not change that; they will still be the Condorcet winner and will still win. If the Black winner was not a Condorcet candidate, they must have been the Borda count winner. Moving them up in the rankings will still leave them as the Borda count winner. And no *other* candidate can become a Condorcet candidate when the original winner moves up in some rankings. So either the original winner becomes the Condorcet candidate and thus wins, or no one is a Condorcet candidate and the original winner still wins by Borda count. Thus Black's method is monotone.

To see the method is Pareto, imagine that A is favored over B by all voters. Then B cannot be the Condorcet candidate because they lose to A head-to-head, and they can't be the Borda count winner because A will have more points than B does. Thus B cannot win.

Finally, we know that Black's method violates independence, because it is Condorcet. We get a concrete example from figure **??**: in the before profile, A is the Condorcet candidate and wins, but in the after profile, we get a three-way tie in which each candidate has three Borda points.

□

Proposition 1.63. *The dictatorship method is monotone, Pareto, and independent, but not Condorcet, anti-Condorcet, or majoritarian.*

Proof. If A is the winner, they are at the top of the dictator's preference list. No improvement on other preference lists can change that, so A remains the dictator's top choice, and thus still wins. So the method is monotone.

If A is higher on ever voter's list than B, then in particular A is above B on the dictator's list and so B does not win. So the dictatorship method is Pareto.

If A is at the top of the dictator's preference list in one profile, and then a second profile has A and B in the same relative position, it can't have B at the top of the dictator's preference list. So B can't win in the second profile. Thus the dictatorship method is independent.

But now consider the following profile, where the third voter is dictator.

A	A	B
B	B	A

Figure 1.9

A is the Condorcet and the majority candidate, while B is the anti-Condorcet candidate. But B wins.

□

Proposition 1.64. *The all-ties method and the monarchy method are monotone and independent, but not Condorcet, anti-Condorcet, majority, or Pareto.*

Proof. These methods are both what we call constant functions, meaning that the output to the method ignores the input completely and doesn't care how anyone votes. These methods are all monotone and independent, because no candidate can win on one profile and lose on another. But they also violate Condorcet, anti-Condorcet, majority, and Pareto, because some profile can have a candidate with a Condorcet candidate, an anti-Condorcet candidate, a majority candidate, or a candidate preferred universally to another candidate—and that has no effect on who wins. □

	anon	neu	unan	dec	maj	Con	AC	mono	Par	ind
Plurality	Y	Y	Y	N	Y	N	N	Y	Y	N
Antiplur	Y	Y	Y	N	N	N	N	Y	N	N
Borda	Y	Y	Y	N	N	N	Y	Y	Y	N
Hare	Y	Y	Y	N	Y	N	N	N	Y	N
Coombs	Y	Y	Y	N	N	N	N	N	Y	N
Copeland	Y	Y	Y	N	Y	Y	Y	Y	Y	N
Black	Y	Y	Y	N	Y	Y	Y	Y	Y	N
Dictator	N	Y	Y	Y	N	N	N	Y	Y	Y
All-ties	Y	Y	N	N	N	N	N	Y	N	Y
Monarchy	Y	N	N	Y	N	N	N	Y	N	Y
Bloc	N	Y	Y	N	N	N	N	Y	Y	N
Agenda	Y	N	Y	N	Y	Y	Y	Y	N	N

Figure 1.10: : A summary of the properties of various social choice methods

1.5 Arrow's Theorem

We've seen that it's impossible to produce a "perfect" social choice function. Fundamentally, the problem is cycles like this, first seen in the proof of proposition **??**:

A	C	B
B	A	C
C	B	A

In this profile, a majority of voters prefer A to B, a majority prefer B to C, and a majority prefer C to A. Even though each individual voter has rational, transitive preferences, the electorate as a whole does not. This is known as the *Condorcet paradox*.

Your textbook presents the following extension that's even worse: suppose we have the preference order

D	C	B
E	D	C
F	E	D
G	F	E
A	G	F
B	A	G
C	B	A

Imagine these represent a series of different policies we could have on some issue, and that we start with policy D. Someone proposes policy C, and that wins by majority vote over D. Then someone else proposes an amendment to B, and that defeats C by majority vote. Then finally, someone else proposes an amendment to shift to A, and that defeats B. In each case the vote was a clear majority vote. But we have moved from D to A, even though every single voter strongly prefers D to A.

We say in proposition **??** that no social choice function can be both independent and Condorcet, due to these loops. But it's worse than that. The insight behind these Condorcet paradox loops lead the economist Kenneth Arrow to formulate the following theorem:

Theorem 1.65 (Arrow)**.** *If a social choice function with at least three candidates satisfies both Pareto and independence, then it must be a dictatorship.*

This is often phrased in the following way:

Theorem 1.66 (Arrow's theorem, alternate version)**.** *It is impossible for a social choice function with at least three candidates to be Pareto, independent, and non-dictatorial.*

Lemma 1.67 (decisiveness lemma)**.** *A social choice function with at least three candidates that satisfies Pareto and independence must be decisive.*

Proof. Suppose we have a profile in which A and B both win. We'll use that assumption to construct a new profile in which we will show *no* candidate can win. But since every voting method has to always produce a winner, that's impossible. This shows that we can't have a profile with two winners.

So suppose we have some profile in which A and B both win. We'll say that x voters prefer A to B, and y voters prefer B to A. Note that neither x nor y can be zero, because if x is zero then all voters prefer B to A, and by the Pareto criterion A can't win.

We can represent our profile like the one in figure **??**. We're technically cheating here because the voters who prefer A to B might have many different specific sets of preferences, so this chart really "should" have a lot more than two columns. But this chart contains all the information we have or need.

x	y
⋮	⋮
A	B
⋮	⋮
B	A
⋮	⋮

Figure 1.11: Profile P in lemma **??**

By our assumption there is at least one more candidate C. Consider another profile Q in which all voters place A, B, and C in third place, with C immediately ahead of B, as in figure **??**:

x	y
A	C
C	B
B	A
⋮	⋮

Figure 1.12: Profile Q in lemma **??**

We claim that C must be the unique winner in profile Q from figure **??**. By the Pareto criterion, B cannot win, since all voters prefer C to B. And between profiles P and Q, no voter has changed the relative rankings of A and B. If A wins and B loses in profile Q, then by independence B cannot win in profile P. Since by hypothesis B wins in profile P, that means A can't win in profile Q.

So neither A nor B can win in profile Q. Further, every other candidate is unanimously beaten by C (and indeed by all three of these candidates), so no other candidate can win. Then C is the only candidate who can possibly win, and since some winner must exist, we know that C is the unique winner in profile Q.

Now we will construct one more profile R. This is just like profile Q, but we interchange B and C.

The relative positions of A and C have not changed between Q and R, so by independence A can't win in R. Then, the relative

x	y
A	B
B	C
C	A
⋮	⋮

Figure 1.13: Profile R in lemma **??**

positions of A and B are the same in P and R; since A wins in P and loses in R, B cannot win in R. But we see that B is unanimously preferred to every candidate except A, so by the Pareto criterion no candidate other than A or B can win.

Thus no candidate can win in profile R. But a voting method has to give a winner for every profile, so this can't happen; that means we can't have a profile like P in which two candidates win simultaneously. That's what we wanted to prove.

□

This lemma leads immediately into another lemma:

Lemma 1.68. *Suppose a social choice function with at least three candidates satisfies Pareto and independence. Suppose there are two profiles in which no voter changes their mind about whether candidate A is preferred to candidate B. If A wins in the first profile, then B cannot win in the second profile.*

Proof. This is a sort of super-independence criterion. It differs from regular independence in that we don't need to assume that "A wins and B loses" in the first profile. We just assume that A wins.

So suppose that A wins in the first profile. By lemma **??**, this method is decisive, so B cannot win.

Then A wins and B loses in the first profile. By independence, B cannot win in the second profile. □

Finally, to prove theorem **??**, we will need define one new operative term. The proof will proceed by showing that for each pair of candidates, there is one specific voter who determines which one must lose.

Definition 1.69. Suppose A and B are two candidates. We say a voter has *dictatorial control* for A over B if, whenever that voter prefers A to B, it will always be the case that B loses.

Note that as written this isn't necessarily symmetric: a voter can have dictatorial control for A over B, but not for B over A.

Lemma 1.70. *If a social choice function with at least three candidates satisfies both Pareto and independence, then for any pair of candidates A and B, there is a voter with dictatorial control for A over B.*

Proof. Fix candidates A and B, and let C be any other candidate. We want to prove that some voter v has dictatorial control for A over B.

Consider a profile in which all voters rank C first, A second, and B third, and then all other candidates after them. By the Pareto criterion, no candidate other than C can win, so C must be the unique winner.

Now imagine the voters, one by one, change their top three places so that they rank A, then B, then C. By the end of this process, A will be the unique winner by the Pareto criterion; and at all points no candidate can win other than A or C, since A will Pareto dominate every candidate other than C. Further by the decisiveness lemma, we can't have A and C both win at the same time. So there is some specific voter who, when they change their mind, causes A to win instead of C. See figure **??**.

X	v	Y
C	C	A
A	A	B
B	B	C
⋮	⋮	⋮

X	v	Y
C	A	A
A	B	B
B	C	C
⋮	⋮	⋮

Figure 1.14: C wins in the first profile and A wins in the second.

We want to show this voter v has dictatorial control for A over B. Imagine any profile P in which voter v ranks A over B. We claim that B must lose. We define a new profile Q by the following algorithm: every voter in X puts C in first, then A, then B in whatever order they ranked them in P. Similarly, every voter in Y ranks A and B first in the same order as in P, and then C third. Voter v ranks A first, then B, then C. We show this schematically in figure **??**

X		v	Y	
C	C	A	A	B
A	B	C	B	A
B	A	B	C	C
⋮	⋮	⋮	⋮	⋮

Figure 1.15: Profile Q from lemma **??**

Now we're ready to make a bunch of independence arguments. In the first profile of figure **??**, candidate C wins, and all candidates give C and B the same relative rankings in profile Q from figure **??**, we know B must lose in profile Q. Since A wins in the second profile of figure **??**, and all voters rank A and C the same as they do in that profile, we know C must lose in profile Q.

By the Pareto criterion, the only possible winners in profile Q are A, B, and C; thus we know A wins in profile Q. But all voters give A and B the same relative ranks in profile Q as in our original profile P. Thus we know B must lose in P by independence.

□

Now we are ready to prove Arrow's theorem.

Proof of theorem **??**. Suppose we have a social choice function that satisfies Pareto and independence. By lemma **??** it must be decisive, and by lemma **??** we know that for any pair of candidates A and B, there is a voter with dictatorial control for A over B.

We first observe that if a voter v has dictatorial control for A over B, no *other* voter w can have dictatorial control for B over A. Suppose we did have such a situation, and then consider a profile where v ranks A first and B second; w ranks B first and A second; and everyone else ranks A and B in the top two spots (for example figure **??**).

v	w	everyone else
A	B	A
B	A	B
⋮	⋮	⋮

Figure 1.16: A profile showing the symmetry of dictatorial control

Since v has dictatorial control for A over B, we know B cannot win. And since w has dictatorial control for B over A, we know A cannot win. But by the Pareto criterion no other candidate can win. That's impossible, so we can't have a voter w

with dictatorial control for B over A.

But we know *some* voter must have dictatorial control for B over A; thus if v has dictatorial control for A over B, they must also have dictatorial control for B over A. So while the definition is not symmetric, under these hypotheses the *relationship* must be.

Thus we have established that for *every* pair of voters A and B, there is a voter with dictatorial control between A and B, such that if they rank A over B then B cannot win, and if they rank B over A then A cannot win. We now claim that the same voter must have dictatorial control over each pair.

So suppose v has dictatorial control between A and B, and w has dictatorial control between B and C. Consider a profile in which v ranks C, then A, then B; w ranks B, then C, then A; and every other voter ranks C then B then A, as in figure **??**.

v	w	everyone else
C	B	C
A	C	B
B	A	A
⋮	⋮	⋮

Figure 1.17: A profile showing every pair must have the same voter with dictatorial control

Then since v prefers A to B, we know B cannot win. Since w prefers B to C, we know C cannot win. And we see that C is

Pareto superior to A and to every other candidate, so neither A nor any other candidate can win. Thus no winner is possible. Since that's a contradiction, we know that two different voters can't have dictatorial control over different pairs.

Thus, we've shown that every pair has a voter with dictatorial control between them. And in fact, this is the same voter for every pair; so there is one voter with the power that, for any two candidates, if they prefer A to B then B cannot win. Thus no candidate other than that voter's first-choice candidate can win the election. Since there must be a winner, that voter's first choice will always win, and our method is dictatorship with that voter being the dictator.

$\square$

Corollary 1.71. *It is impossible for a method to satisfy Pareto, independence, anonymity, and neutrality.*