An introduction to social choice theory, course given at the Nesin Mathematical Village

Daniel Massart

June, 2025

Contents

1 Introduction **2**

2 An example **2**

2.1 First procedure : first past the post 3
2.2 Second procedure : two round majority voting 3
2.3 Third procedure : Condorcet voting 3
2.4 Fourth procedure : Borda voting 4
2.5 Conclusion 4

3 Axiomatic approach : ordinal voting **4**

3.1 Example : dictatorship 5
3.2 Universality 6
3.3 Anonymity 6
3.4 Neutrality 6
3.5 Unanimity 7
3.6 Condorcet criterion 7
3.7 Independance of irrelevant alternatives 7
3.8 Strategy-proofness 8
3.9 Monotonicity 10
3.10 Hierarchy of properties 12
3.11 An impossibility theorem 14

4 Cardinal voting **15**

4.1 Approval voting 16
4.2 General score voting 18
4.3 Majority judgement 18

5 Back to Condorcet : possible fixes **19**
5.1 Unidimensional voter space 19
5.2 Borda+instant runoff = Condorcet 19
5.2.1 Victory margin matrices 19
5.2.2 Instant runoff Borda 20
5.3 Introducing randomness 21
5.4 Ranked pairs 21

1 Introduction

Examples of situations where several (usually many) agents must decide between several alternatives include : the stock market trying to decide the price of some company ; a large network of computers needs to devise a route to send your netflix video to your computer ; attendees at a county fair try to collectively guess the weight of a bull ; your brain (seen as a neuron network) needs to choose between beef, fish, and chicken for dinner ; citizens of a country need to choose their legislators ; then the legislators need to decide on which laws to vote.

Acknowledgements : much of this course is borrowed from my colleagues and friends Xavier Bry and Nicolas Saby, and some of it comes from Rémi Peyre's articles on the french website Images des Maths. See also Hubert Bray's video lecture series, available at

https://professorbray.net/Videos/Democracy/index.php.

2 An example

Let us start with an example. Consider the following situation where 100 voters have to choose one among five alternatives, denoted A, B, C, D, E. The preferences of the voters are recorded in the following array, where the first column indicates the rank in order of preference, and the bottom line indicates the number of voters with the corresponding order of preference.

1	A	B	C	C	D	E
2	B	D	D	E	E	C
3	C	C	B	B	C	B
4	D	E	A	D	B	D
5	E	A	E	A	A	A
	33	16	3	8	18	22

We want to organize a vote that would reflect the collective preference of the voters, or maximize global satisfaction, if you will.

2.1 First procedure : first past the post

(a.k.a. single round majority voting, or plurality rule) A is ranked first by a plurality of voters, so A is elected. Note that for this procedure voters need not give their full order of preference, just their favorite.

But then someone is going to say "look, A is hated by 64% of the population, so how can this be the will of the people ?"

2.2 Second procedure : two round majority voting

(a.k.a. runoff voting) A and E get to the second round, then 64% of the voters like E better than A, so E is elected. Note that the two rounds (or as many rounds as you want) may be done at once if the ballots record the full order of preference of each voter, not just one name. Instant runoff is the 4-round voting variant, where the candidate with the least votes is eliminated in each round.

But then... "66% of the voters like D better than E ! how can..."

2.3 Third procedure : Condorcet voting

(sometimes called majority rule) For each pair of candidates, we look at who is preferred by a majority. The winner is the one who wins every duel : then nobody can object "look, 51% of the voters like somebody else better...".

In this case, we have 5 candidates, so that's $\binom{5}{2} = 10$ duels, C wins all of its 4 duels, so C wins (exercise : prove that the winner is unique !). Note that all duels can be done at once by a computer if the ballots record the full order of preference of each voter.

This voting system looks like a good way to find out what the voters want, once you get past the technicalities of organizing all the duels (very easy with electronic counting, providing the ballots record the full order of preference of each voter). We shall see later that this is not as obvious as it seems, because it could cause a candidate who is hated by 49% of the population to be elected at the expense of a more consensual, if less exciting, candidate. But a more immediate trouble is, the winner need not exist at all ! Consider the following situation :

1	A	C	B
2	B	A	C
3	C	B	A
	33	33	34

A wins against B 66 to 34, B wins against C 67 to 33, C wins against A 67 to 33 : no candidate wins all of his duels. This situation is called a Condorcet paradox, and the triplet (A, B, C) is called a Condorcet cycle.

So what can we do ? of course we could declare this race to be a tie (which is completely ok if you want a ranking of restaurants in İstanbul), and randomly select a winner (where to have dinner tonight). We might be less comfortable with randomness when it comes to selecting a leader for our country, or a restaurant for your wedding party, although it should be noted that the Athenians had no trouble with that.

But what are you going to do in the following case ?

1	A	B	C
2	B	C	D
3	D	D	A
4	C	A	B
	33	33	34

A beats B, B beats C and D, C beats A and D, D beats A. We have several, overlapping, Condorcet cycles : (A, B, C, D), (A, B, C) and (A, B, D). How are you going to decide who is tied with whom, before you can even cast lots ?

We'll come back to randomness later.

2.4 Fourth procedure : Borda voting

Condorcet had a friend/rival called Borda (that was XVIIIth century France), who had another idea : each voter gives each candidate a number of points equal to his ranking. Sum the points of each candidate over all the voters. The candidate with the fewest points, that is, the candidate with the highest average ranking, wins. In our case B wins by a wide margin. Note, however, that there is something arbitrary in our point attribution procedure : why not instead give a number of points equal to the square root of the ranking, to (hopefully ; as an exercise, ask yourself if it really does that) handicap divisive candidates, who are hated by some voters ? We shall see later that Borda voting has other problems.

As it is, France being France, those smart people thought very hard, then the government ignored them (Borda) or sent them to jail (Condorcet), and chose the second to worst method, two-round majority voting.

2.5 Conclusion

Four methods (and there are many more), four different winners : maybe the collective preference is not so easy to define as it seems. Then let's do what mathematicians do : take an axiomatic approach to designing a voting system.

3 Axiomatic approach : ordinal voting

Assume we have a **set of candidates** $\mathcal{C} = \{C_1, \ldots, C_n\}$, and a **set of voters** $\mathcal{V} = \{v_1, \ldots, v_N\}$. Throughout this course I'll assume $n \geq 3$, since everything is more or less trivial when there are only two candidates. Each voter v_i has a **set of preferences**, which we record as a total, strict ordering (whence the word ordinal) P(i) of $\mathcal{C}$, that is, a bijection $\mathcal{C} \longrightarrow \{1, \ldots, n\}$. Denote by $\mathcal{P}$ the set of all such bijections.

This is a very un-realistic hypothesis, especially when there are many candidates (e.g. the restaurants of İstanbul) : you cannot be expected to know all the restaurants and be able to rank them, without ties. I make this hypothesis only because it makes my life simpler.

A **voter profile** is an element of $\mathcal{P}^N$: a record of all preferences of all voters. A **voting system** (sometines called social choice functio, or SCF) is a map

$$\mathcal{F} : \mathcal{P}^N \longrightarrow \mathcal{C}.$$

This is called ordinal voting because preferences are just represented by an ordering of $\mathcal{C}$, as opposed to cardinal voting, which we shall see later.

It is often interesting to consider the limit case when there are infinitely many voters, and instead of counting then, we just record the proportion (as a percentage) of voters with a given order of preference. This amounts to considering $\mathcal{V} = [0, 1]$, $\mathcal{P}^N$ by the set of functions from $[0, 1]$ to $\mathcal{P}$, and replacing counting by integration.

If we want to use randomness, we'll introduce a mysterious set Ω, which represents chance as in your probability class, and consider $\mathcal{F}$ to be defined on $\Omega \times \mathcal{P}^N$.

A variation of this definition, called a social welfare function (SWF for short), is when the range of $\mathcal{F}$ is $\mathcal{P}$ instead of $\mathcal{C}$, that is, you want a full ranking of all candidates. For instance, if you want a restaurant for your wedding party, you need an SCF ; if you want restaurants for the next two years, you need an SWF. Every statement in this course has an SCF and an SWF version, but I'll state only one for the sake of not boring you to death.

You can get an SCF from an SWF by just taking the first-ranked candidate ; and you can get an SWF from an SCF by iteration.

3.1 Example : dictatorship

An SCF $\mathcal{F}$ is called a dictatorship if there exists $v \in \mathcal{V}$, called the dictator, such that $\forall P \in \mathcal{P}^N, \mathcal{F}(P) = P(v)$. A random dictatorship is $\mathcal{F} : \Omega \times \mathcal{P}^N$ such that $\forall \omega \in \Omega, \exists v(\omega) \in \mathcal{V}$, such that $\forall P \in \mathcal{P}^N, \mathcal{F}(\omega, P) = P(v(\omega))$: you first randomly choose a citizen, who then gets to be dictator. This is not the product of the wild imagination of some crazy mathematician : it was used by the Athenians, and it is actually one of the fairest possible systems. So get ready, because you may wake up some day as dictator. I would argue this was the point of the Athenians : people will try to be the best version of themselves if they know they may have to lead their country someday. Whether or not it would work well in practice now, is another question : fairness may not be the only parameter for which we need to optimize.

Nevertheless, if we want the voting system to be accepted by the voters, it needs to be considered fair. Here are some desirable qualities a fair system should have.

3.2 Universality

$\mathcal{F}$ is defined on the whole of $\mathcal{P}^N$: people are free to vote however they want, and it will produce a winner. Note that, strictly speaking, none of the procedures I mentioned, except dictatorships, is universal in this sense, because there may be ties ! However the probability of this happening when there are many (e.g. millions) of voters is ridiculously small, so I shall ignore it. On the other hand, this is a serious problem for Condorcet voting, as you may find out if you solve the following

Exercise 3.1. *Consider Condorcet voting with 3 candidates and so many voters that you may view a voter profile as a vector in $\mathbb{R}^6$ (six is the number of orderings of the three candidates) whose coordinates sum to one (as percentages do). Compute the probability of a Condorcet paradox, assuming uniform distribution of the preferences.*

Remark 3.2. *It is far from obvious how realistic the hypothesis of uniform distribution of the preferences actually is. For instance if the choice is about the room temperature, the candidates are the temperatures, which are naturally ordered ; it is unlikely that any voter will have 1st. 22, 2nd 14, 3rd 18 as order of preferences. We'll come back later to this situation of ordered candidates.*

3.3 Anonymity

All voters are considered equal ; in mathematical terms, $\mathcal{F}$ is invariant under permutations of $\mathcal{V}$. More precisely, the group Σ_N of permutations of acts on $\mathcal{V}$ by permuting the voters, so it acts on $\mathcal{P}^N$ by permuting the columns, if you view a voter profile as a double-entry array whose columns are the voters, and whose rows are the candidates.

Anonymity means that $\forall \sigma \in \Sigma_N, \forall P \in \mathcal{P}^N, \mathcal{F}(\sigma.P) = \mathcal{F}(P)$.

Dictatorships are not anonymous : a permutation σ of the voters replaces the dictatorship of voter v by the dictatorship of $\sigma(v)$. On the other hand, a random dictatorship is anonymous, provided all voters have an equal chance of being drawn.

All the voting systems that operate by counting ballots are anonymous.

3.4 Neutrality

All candidates are considered equal : $\mathcal{F}$ is invariant under permutations of $\mathcal{C}$. More precisely, the group Σ_n of permutations of $\mathcal{C}$ acts on $\mathcal{P}$ by permuting the candidates, so it acts on $\mathcal{P}^N$. Neutrality means that $\forall \sigma \in \Sigma_n, \forall P \in \mathcal{P}^N, \mathcal{F}(\sigma.P) = \mathcal{F}(P)$.

All the examples we have seen so far are neutral because the letters are just labels we put on the candidates ; the preferences of the voters do not change if we permute the labels. An example of a non-neutral voting system would be for instance "if there is a tie, the oldest candidate wins".

Note that neutrality implies that $\mathcal{F}$ is onto (i.e. surjective).

3.5 Unanimity

(a.k.a. Pareto principle) If all voters rank A first in the profile P, then $\mathcal{F}(P) = A$. All the examples we have seen so far are unanimous, including dictatorships (if all voters want A, in particular the dictator wants A).

Now come some more problematic properties.

3.6 Condorcet criterion

If there is a Condorcet winner A for the profile P, then $\mathcal{F}(P) = A$. This sounds reasonable enough (who would want to elect a candidate who could be removed by referendum against some other candidate ?), however none of the examples we have seen meets this requirement, except Condorcet voting itself, which, strictly speaking, is not an SCF, because it may fail to produce a winner.

3.7 Independance of irrelevant alternatives

Let us start with an example. Imagine you are ordering ice-cream, so the candidates are the flavors, and the voters are your neurons. You are given the choice between chocolate and strawberry, and you prefer chocolate. Then you realize there is also vanilla on the menu, so you call back the waiter and say "then I'll have strawberry". Your friends give you a slightly puzzled look : it's because your choice function is not IIA. IIA means that the relative ranking of two candidates should not depend on the presence, or absence, of a third candidate.

Formally, if you have a subset $\mathcal{E} \subset \mathcal{C}$ of candidates, and an ordering $P \in \mathcal{P}$, you get an ordering $P_{\mathcal{E}}$ of $\mathcal{E}$ by restricting P to $\mathcal{E}$ (you just forget all comparisons with candidates not in $\mathcal{E}$).

IIA says that if $A \neq B \in \mathcal{C}$, and $P, P' \in \mathcal{P}^N$ are such that $P_{A,B} = P'_{A,B}$, then (in the SWF case) $\mathcal{F}(P)(A) > \mathcal{F}(P)(B) \iff \mathcal{F}(P')(A) > \mathcal{F}(P')(B)$. In the SCF case, it means that under the hypothesis $P_{A,B} = P'_{A,B}$, it is impossible that $\mathcal{F}(P) = A$ and $\mathcal{F}(P') = B$.

Dictatorships are IIA : assume that for some voter v, for all voter profile

P, $\mathcal{F}(P) = P(v)$. Then

$$\begin{aligned}
\mathcal{F}(P)(A) > \mathcal{F}(P)(B) &\Leftrightarrow P(v)(A) > P(v)(B) \\
&\Leftrightarrow P'(v)(A) > P'(v)(B) \\
&\Leftrightarrow \mathcal{F}(P')(A) > \mathcal{F}(P')(B).
\end{aligned}$$

Single round majority voting is obviously not IIA (case in point : USA, 1992, 2000 ; Taïwan, 2000 ; in the UK it has happened so many times they believe it to be a feature, not a bug, of their system). Neither is two-round majority voting (France, 2002). This is often called the spoiler effect, the spoiler being some candidate who has no chance to win, but may change the result of the election. Borda is not IIA, as shown by the following example :

A	C
C	A
B	B
2	3

A (strawberry) has 8 points, B (vanilla) 15, C (chocolate) 7 : C wins.

A	C
B	A
C	B
2	3

A has 8 points, B 13, C 9, A wins. Note that in both cases the relative rankings of A and C are the same. If B, who has absolutely no chance to win, withdrew from the race, C would win ; but by coming second for some voters in the second race, B modifies the end result. So, if we imagine that A and C belong to opposing political parties, A has an incentive to manipulate the race by getting some little known candidate from the same party as C to run.

On the other hand, Condorcet, if it were a well-defined SCF, would be IIA, because if A wins its duel against B in both P and P', you cannot have $\mathcal{F}(P') = B$.

Exercise 3.3. *Is it true that IIA implies the Condorcet criterion ? how about IIA + unanimity ?*

While IIA is about preventing manipulation on the side of the candidates, the next property is about preventing manipulation on the side of the voters.

3.8 Strategy-proofness

Roughly speaking, an SCF or SWF is called strategy-proof if no voter has an incentive to lie about his preferences. That is, for all voter $v \in \mathcal{V}$, for all $P \in \mathcal{P}^N$, there does not exist $P' \in \mathcal{P}^N$ such that $P(v') = P'(v')$ for all $v' \neq v$, and $\mathcal{F}(P')$ is higher than $\mathcal{F}(P)$ in the order of preference $P(v)$, because otherwise, in the profile P, v would have an incentive to lie and change his preferences to $P'(v)$, to get the more favorable (to v) outcome $\mathcal{F}(P')$.

Note that we are only dealing with individual strategies here, not group strategies.

In the language of game theory, strategy-proofness means that the outcome $\mathcal{F}(P)$ is a Nash equilibrium : no player has an incentive to change his mind, unless someone else changes theirs.

Single round, or two-round, or n-round majority voting, is never strategy-proof : you always have an incentive to vote for the least bad candidate that you think can win, rather than for your true favorite. This breeds cynicism, and may explain why voters lose interest in voting, including in some young democracies like Taïwan or Korea.

Borda is not strategy-proof, for the same reason. This was pointed out by Condorcet to Borda, who replied "my method is meant for honest people". The fact that this sounds so hopelessly naive now shows how cynical decades of strategic voting have made us.

Dictatorships are strategy-proof, because the dictator has no reason to lie about his preferences.

Lemma 3.4. *Condorcet is strategy proof.*

Proof. The idea is that if you like A better than B, lying about that by saying you like B better, is never going to improve A's chances in any duel.

Formally, consider $P, P' \in \mathcal{P}^N$ such that $P(v') = P'(v')$ for all $v' \neq v$, and assume

$$P(v)(\mathcal{F}(P')) > P(v)(\mathcal{F}(P)) \tag{1}$$

Since $\mathcal{F}(P)$ is the Condorcet winner for the profile P, $\mathcal{F}(P)$ wins its duel agains $\mathcal{F}(P')$ in the profile P. By (1), this duel is won thanks to voters other than v. But since $P(v') = P'(v')$ for all $v' \neq v$, this means that $\mathcal{F}(P)$ also wins its duel agains $\mathcal{F}(P')$ in the profile P', so $\mathcal{F}(P')$ cannot be the Condorcet winner for the profile P'. $\square$

Why strategy-proofness is desirable in the first place, is not completely obvious, although we shall see some examples where it clearly is (and some where it is not that clear-cut). Broadly speaking it could be said that strategic voting displaces the focus of the voters, from looking for the common good, to optimizing for one's objectives, making voting, in the words of ??, "more of a game of skill than a test of voters' preferences".

In the context of, for instance, attendees at a county fair trying to guess the weight of a bull, it is important that each voter gives his honest opinion independantly of others. More generally, whenever the question being debated has an objective answer, it is best to rule out strategic voting.

Here is an example showing how the possibility of strategic voting with Borda may lead to disaster : assume four candidates are running for president, A and B are your run-of-the-mill politicians, but C and D are fringe candidates who want to make an alliance with Sauron and plunge Middle Earth into darkness, D being the craziest. Almost nobody supports them, so the sincere voter profile looks like

A	B
B	A
C	C
D	D
60	40

A is the Borda winner (and also the Condorcet winner). Now, B has a bright idea : he quietly lets his supporters know that they should place C and D ahead of A on their ballots, so the profile becomes

A	B
B	C
C	D
D	A
60	40

and now, thanks to this brilliant strategy, B wins. Of course A supporters are not so easily fooled, so they use the same tactics :

A	B
C	C
D	D
B	A
60	40

and C wins, and all hell breaks loose. Crazy as it seems, this situation is a Nash equilibrium : nobody has an incentive to change their mind, until maybe Gandalf steps in and forces A and B supporters to the negotiating table (although he would do an even better job changing the voting system). That could have been avoided by giving the voters an incentive to vote according to their true preferences (assuming, of course, there are sufficiently many voters so statistics ensure the crazy ones are a tiny minority).

Another intuitive-sounding yet hard-to-express property, which will be useful to us from technical reasons :

3.9 Monotonicity

Informally, $\mathcal{F}$ is said to be monotonic if modifying P by raising the ranking of A in some P(v) cannot prevent A from being elected (SCF version), or cannot lower the ranking of A in $\mathcal{F}(P)$ (SWF version).

Formally, an SWF $\mathcal{F}$ is said to be monotonic if $\forall P, P' \in \mathcal{P}^N, \forall A, B \in \mathcal{C}$, if $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$ (i.e. A is collectively preferred to B in the profile P), and

$$\forall v \in \mathcal{V}, P(v)(A) > P(v)(B) \Rightarrow P'(v)(A) > P'(v)(B)$$

(i.e. support among voters for A relative to B only increases when going from P to P'), then $\mathcal{F}(P')(A) > \mathcal{F}(P')(B)$ (i.e. A is still collectively preferred to B in the profile P').

Likewise, an SCF $\mathcal{F}$ is said to be monotonic if $\forall P, P' \in \mathcal{P}^N, \forall A, B \in \mathcal{C}$, if $\mathcal{F}(P) = A$ (i.e. A wins the election), and

$$\forall v \in \mathcal{V}, P(v)(A) > P(v)(B) \Rightarrow P'(v)(A) > P'(v)(B),$$

then $\mathcal{F}(P') \neq B$.

Remark 3.5. *Beware that monotonicity is often defined as "if A wins an election, and the same election is run again, and more people vote for A (meaning the only change is that some people who didn't vote for A, do), then A should win again". With this definition it is obvious that first past the post is monotonic, although two-round majority voting is not. However, this is not exactly what we mean by monotonic, as shown by our next example.*

Here is an example of a voter profile P with three candidates and 21 voters, in which, under the first-past-the-post system, candidate A wins (although 13 voters out of 21 like B better than A):

A	B	C
B	A	B
C	C	A
8	7	6

Now consider the modified profile P', where two voters change from the third column to the second :

A	B	C
B	A	B
C	C	A
8	9	5

Popular support for B relative to A hasn't changed : the voters who liked A better than B, still do, that is, $\forall v \in \mathcal{V}, P(v)(A) > P(v)(B) \Rightarrow P'(v)(A) > P'(v)(B)$, but now B wins the election (then again, A was not a particularly legitimate winner).

Here is an example showing why two-round majority voting is not monotonic : you may hurt your favorite's chances by voting for them !

1	C	A	B
2	A	C	A
3	B	B	C
	8	6	7

C and B get to the second round, C wins because voters who rank A first rank C second. C is collectively preferred to A, although only 8 voters out of 21 prefer C to A. Now consider the modified profile, where two voters change from the third column to the first :

1	C	A	B
2	A	C	A
3	B	B	C
	10	6	5

Global support for C has definitely increased, but now A gets to the second round and beats C 11 to 10. Note that this example also shows that two-round majority voting is not strategy-proof : the two voters who switched had an incentive to lie by sticking to their former preferences. This is not a coincidence as we shall see next.

3.10 Hierarchy of properties

If you are confused about the relationship between all those complicated definitions, here is some clarification :

Lemma 3.6. *Monotonic $\Rightarrow$ IIA.*

Proof. (SCF version) IIA may be expressed as follows : $\forall P, P' \in \mathcal{P}^N, \forall A, B \in \mathcal{C}$, the equivalence

$$\forall v \in \mathcal{V}, P(v)(A) > P(v)(B) \Leftrightarrow P'(v)(A) > P'(v)(B)$$

(which is just a fancy way of saying $P_{A,B} = P'_{A,B}$) entails

$$\mathcal{F}(P) = A \Leftrightarrow \mathcal{F}(P') = A.$$

Applying monotonicity in both directions (from P to P', then back), you get IIA. □

Lemma 3.7. *Strategy-proof $\Rightarrow$ monotonic.*

Proof. (SWF version) assume $\mathcal{F}$ is strategy-proof. Take $P, P' \in \mathcal{P}^N$ and $A, B \in \mathcal{C}$, such that $\forall v \in \mathcal{V}, P(v)(A) > P(v)(B) \Rightarrow P'(v)(A) > P'(v)(B)$, and $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$. We need to prove that $\mathcal{F}(P')(A) > \mathcal{F}(P')(B)$.

Case 1 : $P(v_i) = P'(v_i) \forall i \neq 1$, i.e. the profiles P and P' are only different in the first column.

Case 1.1 : $P(v_1)(A) > P(v_1)(B)$, i.e. the voter v_1 prefers A to B in the profile P, hence also in P'. Therefore, if we had $\mathcal{F}(P')(A) < \mathcal{F}(P')(B)$, then in the profile P', the voter v_1 (who likes A better than B) would have an incentive to lie by changing his preferences to $P(v_1)$. This contradicts strategy-proofness.

Case 1.2 : $P(v_1)(A) < P(v_1)(B)$. Then v_1 would have an incentive to lie in the profile P.

Case 2 (general case) : we build a sequence of profiles $P = P_0, P_1, \ldots, P_N = P'$, such that P_{i+1} and P_i only differ in the i-th column. Applying the first case N times, we get $\mathcal{F}(P')(A) > \mathcal{F}(P')(B)$. □

So monotonicity is an intermediate property between the stronger strategy proofness and the weaker IIA.

Lemma 3.8. *Strategy-proof + neutral ⇒ unanimous.*

Proof. (SWF version) assume that for some $P \in \mathcal{P}^N$, $\forall v \in \mathcal{V}$, $P(v)(A) > P(v)(B)$, and let us prove that $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$. By IIA, which follows from strategy-proofness, it is enough to work with $P_{A,B}$, which is just

A	...	A
B	...	B
v_1	...	v_N

Assume $\mathcal{F}(P)(A) < \mathcal{F}(P)(B)$. Then by neutrality, if we reverse the parts of A and B in P, to get the profile P' such that $P'_{A,B}$ is

B	...	B
A	...	A
v_1	...	v_N

we have $\mathcal{F}(P')(A) > \mathcal{F}(P')(B)$. Let us consider the sequence of profiles $P = P_0, P_1, \ldots, P_N = P'$, such that P_{i+1} and P_i only differ by swapping A and B in the i-th column. Since we begin with $\mathcal{F}(P)(A) < \mathcal{F}(P)(B)$ and end with $\mathcal{F}(P')(A) > \mathcal{F}(P')(B)$, there is an

$$i_0 = \min \{ i : \mathcal{F}(P_i)(A) > \mathcal{F}(P_i)(B) \} .$$

Thus we have

$$\begin{gathered} P_{i_0}(v_{i_0})(B) > P_{i_0}(v_{i_0})(A) \\ \mathcal{F}(P_{i_0})(A) > \mathcal{F}(P_{i_0})(B) \\ P_{i_0-1}(v_{i_0})(B) < P_{i_0-1}(v_{i_0})(A) \\ \mathcal{F}(P_{i_0-1})(A) < \mathcal{F}(P_{i_0-1})(B) \end{gathered}$$

Therefore, in the profile P_{i_0}, the voter v_{i_0}, who prefers B, can get B ahead of A (which is a better outcome for v_{i_0}) by claiming to prefer A. This contradicts strategy-proofness. □

And now, the bad news.

3.11 An impossibility theorem

Theorem 3.9. *Assume an SWF $\mathcal{F}$ is unanimous, neutral and monotonic. Then $\mathcal{F}$ is a dictatorship.*

Corollary 3.10. *Assume an SWF $\mathcal{F}$ is strategy-proof and neutral. Then $\mathcal{F}$ is a dictatorship.*

Remark 3.11. *This is a weak version of two classical results in the subject : Arrow's theorem (1951), which only assumes IIA and unanimity, and Gibbard-Satterthwaite's theorem (1975), which assumes strategy-proofness but not neutrality. The simple and elegant proof that follows is due to Terence Tao.*

Proof. (of Theorem 3.9). Define a quorum (or deciding coalition) to be a subset $\mathcal{S}$ of $\mathcal{V}$ such that $\forall P \in \mathcal{P}^N$, $\forall A, B \in \mathcal{C}$,

$$(\forall v \in \mathcal{S}, P(v)(A) > P(v)(B)) \Longrightarrow (\mathcal{F}(P)(A) > \mathcal{F}(P)(B))\,.$$

Quorums exist : for instance, $\mathcal{V}$ is a quorum, by unanimity. More interestingly,

Lemma 3.12. *For all $v \in \mathcal{V}$, either $\mathcal{V} \setminus \{v\}$ is a quorum, or v is a dictator.*

Proof. (of Lemma 3.12). Assume $\mathcal{V} \setminus \{v\}$ is not a quorum :

$$\exists P_0 \in \mathcal{P}^N, \exists A_0, B_0 \in \mathcal{C}, \forall v' \neq v, P_0(v')(A_0) > P_0(v')(B_0),$$

but $\mathcal{F}(P_0)(A_0) < \mathcal{F}(P_0)(B_0)$. Observe that $P_0(v)(A_0) < P_0(v)(B_0)$, otherwise, by unanimity, $\mathcal{F}(P_0)(A_0) > \mathcal{F}(P_0)(B_0)$.

Let us prove that v is a dictator, that is,

$$\forall P \in \mathcal{P}^N, \forall A, B \in \mathcal{C}, P(v)(A) < P(v)(B) \Rightarrow \mathcal{F}(P)(A) < \mathcal{F}(P)(B).$$

So take $P \in \mathcal{P}^N$, and $A, B \in \mathcal{C}$ such that $P(v)(A) < P(v)(B)$, and let us prove that $\mathcal{F}(P)(A) < \mathcal{F}(P)(B)$. By neutrality, we may assume $A = A_0, B = B_0$. Then,

$$\forall v' \in \mathcal{V}, P_0(v')(A) < P_0(v')(B) \Rightarrow P(v')(A) < P(v')(B),$$

that is, B is at least as well ranked, relative to A by the voters in P, as in P_0, which is not difficult since in P_0 everybody but v prefers A to B.

Therefore, by monotonicity, and because $\mathcal{F}(P_0)(A_0) < \mathcal{F}(P_0)(B_0)$, we have $\mathcal{F}(P)(A_0) < \mathcal{F}(P)(B_0)$, so v is a dictator. ☐

Lemma 3.13. *If S and T are quorums, then so is $S \cap T$.*

Proof. (of Lemma 3.13) Assume that for some $P \in \mathcal{P}^N$ and $A, B \in \mathcal{C}$, $\forall v \in S \cap T$, we have $P(v)(A) > P(v)(B)$, and let us prove that $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$.

Furthermore, we may assume $\forall v \notin S \cap T$, we have $P(v)(A) < P(v)(B)$, because if $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$ holds in this case, by monotonicity, it holds also in the general case.

By IIA we may assume that there exists $C \in \mathcal{C}$ such that $P_{A,B,C}$ is

B	A	C
A	C	B
C	B	A
$S \setminus T$	$S \cap T$	$T \setminus S$

Then $\forall v \in S$, $P(v)(A) > P(v)(C)$, so, S being a quorum, $\mathcal{F}(P)(A) > \mathcal{F}(P)(C)$. Likewise, $\forall v \in T$, $P(v)(C) > P(v)(B)$, so, T being a quorum, $\mathcal{F}(P)(C) > \mathcal{F}(P)(B)$. Therefore $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$. □

Not let us finish the proof of Theorem 3.9 : if there is no dictator, then for all $v \in \mathcal{V}$, $\mathcal{V} \setminus \{v\}$ is a quorum, and since

$$\bigcap_{v \in \mathcal{V}} \mathcal{V} \setminus \{v\} = \emptyset,$$

we get that $\emptyset$ is a quorum, which contradicts unanimity. □

4 Cardinal voting

Since it seems there is no optimal ordinal voting method, why not try something else : instead of taking the range of $\mathcal{F}$ to be $\mathcal{C}$ (SCF) or $\mathcal{P}$ (SWF), why not take it to be (some subset of) $\mathbb{R}$, each voter giving each candidate a grade, and then somehow aggregating those grades to get a collective grade for each candidate ? The first aggregation method that comes to mind is just taking the sum of all individual scores (note that taking the mean would yield the same result) : for each voter v, $P(v)$ is now a map $\mathcal{C} \to \mathbb{R}$, and

$$\begin{aligned} \mathcal{F} : \mathcal{C} &\longrightarrow \mathbb{R} \\ A &\longmapsto \sum_{i=1}^{N} P(v_i)(A). \end{aligned}$$

This method, like our previous ones, is not universal strictly speaking, because there may be ties ; again, we'll ignore this problem by restricting the domain of $\mathcal{F}$ to the domain where ties do not happen. It is anonymous (because a score point is a score point whomever it comes from) and neutral (a point is a point whomever it is given to). It is unanimous because if everybody gives a higher score to candidate A than to candidate B, then the aggregate of A is higher than that of B.

For cardinal voting we have to adapt a little bit the definition of monotonicity, because candidates may be tied in some $P(v)$, so we shall say $\mathcal{F}$ is monotonic if $\forall P, P' \in \mathcal{P}^N$ and $A, B \in \mathcal{C}$, such that

- $\forall v \in \mathcal{V}, P(v)(A) > P(v)(B) \Rightarrow P'(v)(A) > P'(v)(B)$,
- $\forall v \in \mathcal{V}, P'(v)(A) < P'(v)(B) \Rightarrow P(v)(A) < P(v)(B)$, (this is the new condition required by the possibility of ties)
- $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$,

we have $\mathcal{F}(P')(A) > \mathcal{F}(P')(B)$.

The definition of IIA also changes a little bit : we say $\mathcal{F}$ is IIA if $\forall P, P' \in \mathcal{P}^N, \forall A, B \in \mathcal{C}$, the equivalences

$$\begin{gathered}\forall v \in \mathcal{V}, P(v)(A) > P(v)(B) \Leftrightarrow P'(v)(A) > P'(v)(B) \\ \text{and } P'(v)(A) < P'(v)(B) \Leftrightarrow P(v)(A) < P(v)(B)\end{gathered}$$

entail

$$\mathcal{F}(P)(A) > \mathcal{F}(P)(B) \Leftrightarrow \mathcal{F}(P')(A) > \mathcal{F}(P')(B).$$

4.1 Approval voting

This is the simplest possible (non-trivial) cardinal voting method : the range of $\mathcal{F}$ is $\{0, 1\}$ (0 for disapprove, 1 for approve), and the aggregation is the sum, that is, each voter's ballot is a map $P(v) : \mathcal{C} \to \{0, 1\}$, and the aggregation is

$$\begin{aligned}\mathcal{F} : \mathcal{C} &\longrightarrow \mathbb{N} \\ A &\longmapsto \sum_{i=1}^{N} P(v_i)(A).\end{aligned}$$

In short, 1 stands for approval from any voter, and the candidate with most approvals wins. You are already familiar with this method : you have been using it ever since you've been giving likes or dislikes to videos, or using doodle to determine when to go out with your friends.

The Spartans used an amusing variant (see Plutarch's Life of Illustrious Men), where instead of counting ballots, they recorded how loud the assembly of voters shouted for a given candidate. I don't know whether voters were allowed to shout at the tops of their lungs or not at all (in which case, it was indeed approval voting) or if they could modulate the loudness of their shout (in which case it was more general score voting), and in any case it was not anonymous because some people are louder than others, but voters could shout for as many candidates as they wished, so at least it was definitely not uninominal voting : Sparta 1, modern democracies 0.

Lemma 4.1. *Approval voting meets the Condorcet criterion.*

Proof. Assume A wins their duel against B, that is,

$$|\{v \in \mathcal{V} : P(v)(A) = 1, P(v)(B) = 0\}| > |\{v \in \mathcal{V} : P(v)(B) = 1, P(v)(A) = 0\}|$$

then $\sum_{i=1}^{N} P(v_i)(A) > \sum_{i=1}^{N} P(v_i)(B)$, so $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$. □

Lemma 4.2. *Approval voting is monotonic (hence IIA).*

Proof. Take $P, P' \in \mathcal{P}^N$ and $A, B \in \mathcal{C}$, such that

1. $\forall v \in \mathcal{V}, P(v)(A) > P(v)(B) \Rightarrow P'(v)(A) > P'(v)(B)$
2. $\forall v \in \mathcal{V}, P'(v)(A) < P'(v)(B) \Rightarrow P(v)(A) < P(v)(B)$
3. $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$.

We need to prove that $\mathcal{F}(P')(A) > \mathcal{F}(P')(B)$.

But $\mathcal{F}(P)(A) > \mathcal{F}(P)(B)$ means

$$\sum_{i=1}^{N} P(v_i)(A) > \sum_{i=1}^{N} P(v_i)(B),$$

or equivalently,

$$|\{v \in \mathcal{V} : P(v)(A) = 1, P(v)(B) = 0\}| > |\{v \in \mathcal{V} : P(v)(B) = 1, P(v)(A) = 0\}|$$

and Conditions 1. and 2. above mean

$$\{v \in \mathcal{V} : P(v)(A) = 1, P(v)(B) = 0\} \subset \left\{v \in \mathcal{V} : P'(v)(A) = 1, P'(v)(B) = 0\right\}$$
$$\{v \in \mathcal{V} : P(v)(B) = 1, P(v)(A) = 0\} \supset \left\{v \in \mathcal{V} : P'(v)(B) = 1, P'(v)(A) = 0\right\}$$

so

$$\sum_{i=1}^{N} P'(v_i)(A) - \sum_{i=1}^{N} P'(v_i)(B) > \sum_{i=1}^{N} P(v_i)(A) - \sum_{i=1}^{N} P(v_i)(B) > 0$$

whence $\mathcal{F}(P')(A) > \mathcal{F}(P')(B)$. □

However, approval voting is not strategy-proof, as shown by the following example : suppose A and B are tied, and you think both are ok, but you slightly prefer A. Then you have an incentive to break the tie by giving 0, instead of 1, to B. I would argue this is a relatively harmless kind of manipulation. At least you never have an incentive not to vote for your favorite candidate.

It is amazing that this method, which is extremely easy to implement, already widely used in almost all domains of life, and vastly superior to all methods currently used in politics, is not the preferred method for electing lawmakers and presidents.

4.2 General score voting

You could argue that the aforementioned manipulation of approval voting is not really a manipulation, it's just that approval voting doesn't give you the possibility of finely expressing your preference, so why not extend the range of $P(v)$? It is probably impractical to take it to be $\mathbb{R}$, but why not follow the practice of e-commerce and use stars, from 1 up to 5 ? This method is again monotonic, for the same reason as approval voting, but it is not Condorcet, as shown by the following example, with only two candidates A and B, and 10 voters.

Assume 6 voters give A 3 stars and B 2, while 4 voters give B 5 and A 1. Then A is the Condorcet winner, but B wins by 32 stars to 22. In this case it is not clear whether the Condorcet criterion should prevail : after all, B does seem to maximize voter satisfaction. The problem is, what if the 4 voters with rather extreme opinions were not sincere ? Imagine they actually thought A deserved 2 stars and B 3, but they didn't want B to lose, so they voted strategically. When the other 6 voters realize it, they will start using the same strategy, so everybody ends up giving either 5 or 1 star, which takes us back to approval voting.

So, is there a way to keep the upside of approval voting, while still giving voters more possibility for nuanced opinions ?

4.3 Majority judgement

I would argue that the problem in the previous method is that the mean gives very skewed information. You know what the mean income of top mathematicians (defined by having at least 3 papers published in the Annals of Mathematics) is ? it is roughly 10 000 000 \$/year. If that seems a lot to you, the reason is that the definition I gave of a top mathematician is carefully tailored to include Jim Simons, and as few other people as possible. The mean is exceedingly sensitive to outliers. So, the idea is to use the median instead. The median income of top mathematicians is probably closer to 100 000 \$/year, and that's only because many of them find employment in high-income countries such as Switzerland, Singapore, the UAE, or the USA (a good example of voting with their feet). Don't say you had no warning if you choose mathematics over the bank.

If $x_1, \ldots, x_N$ are real numbers, the median interval of $x_1, \ldots, x_N$ is the set M of all real numbers such that for all $z \in M$,

$$|\{i = 1, \ldots, N : x_i > z\}| = |\{i = 1, \ldots, N : x_i < z\}|$$

(exercise : prove it is actually an interval, possibly a point). The median $\mathrm{med}(x_1, \ldots, x_N)$ of $x_1, \ldots, x_N$ is the middle point of the median interval.

So the idea of median voting is that for any candidate C,

$$\mathcal{F}(C) = \mathrm{med}(P(v)(C) : v \in \mathcal{V}).$$

5 Back to Condorcet : possible fixes

The next three methods are strictly for nerds, so let's imagine we are all citizens of Nerdişehir, and we want to elect our mayor.

5.1 Unidimensional voter space

5.2 Borda+instant runoff = Condorcet

Let us come back to Borda, the forgotten man of this story, for a minute. First let us define the victory margin matrix.

5.2.1 Victory margin matrices

Assume for example that we have four candidates A, B, C, D, and for some voter v, $P(v)$ is $A > B > C > D$. We may define a victory margin matrix $MV(P, v)$ for the voter v, as

	A	B	C	D
A	0	1	1	1
B	-1	0	1	1
C	-1	-1	0	1
D	-1	-1	-1	0

If the preferences of v are different, the matrix $MV(P, v)$ is still an antisymmetric matrix all of whose non-diagonal entries are ± 1, the entry A, B for instance is 1 if v likes A better than B, -1 otherwise. Then the total victory margin matrix is

$$MV(P) = \sum_{v \in \mathcal{V}} MV(P, v).$$

so each voter v contributes 1 to the A, B entry of $MV(P)$ if v likes A better than B, -1 otherwise.

In general, for $P \in \mathcal{P}^N$, $A, B \in \mathcal{C}$, we define the margin of victory of A over B as

$$mv(P)(A, B) = |\{v \in \mathcal{V} : P(v)(A) > P(v)(B)\}| - |\{v \in \mathcal{V} : P(v)(A) < P(v)(B)\}| .$$

We then define the victory margin matrix $MV(P)$ as the square $n \times n$ antisymmetric matrix whose i, j-th entry is $mv(P)(C_i, C_j)$.

A nice feature of the particular of the victory margin matrix is that it makes it easy to see wheteher or not there is a Condorcet winner : it is (if it exists) the only row, all of whose non-diagonal entries are positive.

The victory margin matrix does not determine P, because you could for instance permute A and B in the order of preferences $P(v)$ of some voter v, and permute them again in some other $P(v')$, then the matrix $MV(P)$

hasn't changed, but P has. However we shall see that $MV(P)$ determines the Borda winner.

In the 4-candidate example above, Candidate A, for instance, is the favorite of Voter v, then the A-row of of $MV(P,v)$ contributes 3 to the sum of the entries of the A-row of $MV(P)$. The second favorite B contributes 1 to the sum of the B-row, the 3rd favorite C contributes -1 to the C-row sum, and the last favorite D contributes -3 to the D-row sum.

So the A-row sum $\Sigma_A(P)$ of $MV(P)$ is $3\times$ the number $n(A,1)$ of voters who rank A 1st, $+1\times$ the number $n(A,2)$ of voters who rank A 2nd, $-1\times$ the number $n(A,3)$ of voters who rank A 3rd, $-3\times$ the number $n(A,4)$ of voters who rank A 4th, that is,

$$\Sigma_A(P) = 3n(A,1) + n(A,2) - n(A,3) - 3n(A,4).$$

Remember the Borda score $B_A(P)$ of A is is $1\times$ the number $n(A,1)$ of voters who rank A 1st, $+2\times$ the number of voters who rank A 2nd, $+3\times$ the number of voters who rank A 3rd, $+4\times$ the number of voters who rank A 4th, that is,

$$B_A(P) = n(A,1) + 2n(A,2) + 3n(A,3) + 4n(A,4).$$

Observe that

$$\Sigma_A(P) + 2B_A(P) = 5(n(A,1) + n(A,2) + n(A,3) + n(A,4)) = 5N,$$

so from $MV(P)$ we can deduce the Borda ranking.

If we have n candidates instead of four, each time A is ranked 1st, contributes $n-1$ to $\Sigma_A(P)$, and each time A is ranked 1st, contributes $n-2k+1$ to $\Sigma_A(P)$.

Therefore we have

$$\begin{aligned}
\Sigma_A(P) &= \sum_{k=1}^{n}(n-2k+1)n(A,k) \\
B_A(P) &= \sum_{k=1}^{n} kn(A,k) \\
\Sigma_A(P) + 2B_A(P) &= \sum_{k=1}^{n}(n+1)n(A,k) = (n+1)N.
\end{aligned}$$

5.2.2 Instant runoff Borda

We have seen that Borda is not a very good method, but what if we combine it with instant runoff, that is, at each of the n rounds, the Borda loser (that is, the candidate with the lowest row sum) is eliminated, and then we do a

Borda vote on the remaining candidates ? The good thing is we don't have to redo all the calculations, now that we have the victory margin matrix, if for instance D is the loser, we just remove the D-row and the D-column from the matrix, and recompute the row sums.

And now the cool thing is that it is very easy that this method is Condorcet. Note that the sum of the entries of an antisymmetric matrix is zero, so the sum, over all rows, of the row sums, is zero. So the lowest row sum is always negative. Therefore, the Condorcet winner, whose row sum is positive, and remains positive after the matrix is reduced, is never eliminated.

Apart from this remarkable property, Borda instant runoff is not a particularly good method, as shown by the following example, but the proof of the Condorcet criterion is so cool I had to show it.

5.3 Introducing randomness

5.4 Ranked pairs