The mathematics of voting, power, and sharing - Part 1

Voting systems

A voting system or a voting scheme is a way for a group of people to select one from among several possibilities.

If there are only two alternatives between which to choose, then it is easy: the alternative that is preferred by the majority wins.

If there is only one person doing the choosing, then things are easy again: you know your own mind, and you know you prefer A to B, and B to C, and C to D. And if by any chance alternative B were to disappear, then you still prefer A to C and C to D.

It is when several people have to choose among more than two alternatives that things become trickier.

The following is a simple example that illustrates one of the oldest voting paradoxes. Suppose a sizable group of people (say 60) will meet for a celebration in a restaurant, and the restaurant manager wants them to pick one menu for the whole group. For the main courses, she initially offers the choice between salmon and chicken. The organizers consult their group, and find that a majority prefers salmon. But when they call up the restaurant owner, she apologizes: her fish supplier has become less reliable, and she is now offering a choice between chicken and beef instead. After the group is consulted, it turns out that a majority now prefers the chicken choice. In summary, the group

- prefers salmon over chicken;
- prefers chicken over beef.

A day later, the restaurant manager calls back; she has switched to another supplier, who has a superb fish assortment, and she can again offer salmon. On the other hand, there is a salmonella scare and many of her patrons are shying away from chicken. So the choice is now between salmon and beef. The organizers feel sure, in view of the ranking above, that their group will largely prefer salmon, but when they ask, they find a clear majority for beef: the group

- prefers beef over salmon.

"Oh well," they think, "people are fickle, and some of them must have changed their minds." Yet, this was not the case: every single person polled had a clear ranking of the three possibilities and stuck to that ranking in a consistent way. Nonetheless, even though every single individual is entirely consistent, the group is not. Let's look at a numerical example: suppose that

- 25 people rank
  1. salmon
  2. chicken
  3. beef
- 20 people rank
  1. chicken
  2. beef
  3. salmon
- 15 people rank
  1. beef
  2. salmon
  3. chicken

So, all in all, 40 people rank salmon higher than chicken, another group of 45 ranks chicken higher than beef, and yet a different grouping of 35 ranks beef higher than salmon: the paradoxical behavior of the group is explained.

This kind of paradox in fact happens all the time, and for things far more serious than menus for celebration dinners such as presidential elections or votes in congress.

In the case where this type of paradox doesn't happen, that is, when there is one alternative that is always preferred by a majority (although not always the same majority) if it were pitted in a one-on-one race against any one of the others, then we call the winning alternative the "Condorcet" winner [this would be the case for the "chicken" choice in the example above if the third group had changed their ordering to 1) beef, 2) chicken, 3) salmon].

We have just seen that there doesn't always exist a Condorcet winner. But when there exists one, it seems fair that that should be the winning choice for the whole group. Or does it?

There are many different systems in existence to select the "winner." Because the Condorcet method doesn't always yield a winner, it is not used a lot. Other methods are:

- **Plurality**: The candidate who is ranked in first place most often, wins. This is the way in which senators or members of congress are elected in the US in every state.
- **Plurality with run-off**: The two candidates with the most first places are retained, and then a second round run-off election is held between them. This is the system used in the election of the president of France.
- **Sequential run-off or the Hare system**: First, the candidate with the fewest first places is removed, then (after his/her votes have been redistributed among the remaining candidates) the next-bottom candidate, and so on .... This system has been used for years in Australia, Ireland, and also in Cambridge, Mass., and in New York City (although not in situations where only one winner has to be selected, but where several seats are available). It was also recently adopted in several local elections in California
- **Borda count**: If there are N candidates, then every voter gives N points to his/her first, $N-1$ to the second choice, ... The points that all the voters gave are then added, and the candidate with the most points wins. This system is often used in clubs to decide on admission (or not) of new members.

The video illustrates an example where different methods can lead to different outcomes ...

Let's illustrate these paradoxical situations with a few more examples.

Example: Paradox with run-off or sequential run-off. For her assignment in *Math Alive*, a student asks 17 of her friends what kind of breakfast they prefer. Here are their answers:

number of people for each ranking	**6**	**5**	**4**	**2**
cereal	1	2	3	2
danish	2	3	1	1
bagel	3	1	2	3

Let's first get rid of the alternative that got fewest first places: bagel (which had (fill in the blanks) ___, while danish had ___, cereal had ___). That leaves cereal and danish.

With only these two alternatives remaining, the preferences are:

	6	**5**	**4**	**2**
cereal				
danish				

so that ________ wins, because it has the most first places now.

But if the last group of **2** votes changes it mind, and decides to rank cereal above danish instead of the other way around, what happens then? Surely cereal's chances of winning must be better now?! Let's check:

	6	**5**	**4**	**2**
cereal	1	2	3	1
danish	2	3	1	2
bagel	3	1	2	3

The item with fewest first places is now danish (___ versus ___ for cereal, ___ for bagel). Reassigning danish's votes, we get:

	6	**5**	**4**	**2**
cereal				
bagel				

so that ________ wins, and ________ loses, even though more voters preferred cereal than before ... (one could imagine the following newspaper headline after the Irish elections, which use this "Hare" system: O'Grady loses seat, although he would have won if fewer people had voted for him ...).

Other example: Paradox with Borda scheme: A club of 25 people are planning an outing. They have narrowed down the choices to a trip to the beach, a hike in the mountains, or a day in NYC. Their preference schedule is the following:

	13	**10**	**2**
beach	2	1	3
mountains	3	2	1
NYC	1	3	2

This is in fact a case where there is a Condorcet winner: in the one-on-one contests NYC always wins:

- beach vs. NYC: ___ people prefer NYC; ___ people prefer beach;
- mountains vs. NYC: ___ people prefer NYC; ___ people prefer mountains.

NYC also wins the plurality vote, and is also the winner under the run-off scheme. In a Borda-count, we find the following totals of points:

beach:	___ $\times 3$	+	___ $\times 2$	+	___ $\times 1 =$	___
mountains:	___ $\times 3$	+	___ $\times 2$	+	___ $\times 1 =$	___
NYC:	___ $\times 3$	+	___ $\times 2$	+	___ $\times 1 =$	___

This does not lead to the same winner, even though NYC won by several other methods .... Many of these paradoxes have been known for a very long time. For a long time also, many people have tried to think of schemes that would avoid such paradoxes. Until, in the early 1950's, Kenneth Arrow attacked the problem in a mathematical way.

He started by listing properties that a "fair" voting scheme should have. Imagine that you have two voters, 1 and 2, and three issues, A, B, and C. A voting scheme is a way of distilling out of the two individual preference schemes a ranking for the "group." We shall say that

$A\ R_1\ B$	if	voter 1 ranks A higher than or equal to B (R for rank) (we allow ties here);
$A\ I_1\ B$	if	voter 1 ranks A and B at the same level (I for indifferent);
$A\ P_1\ B$	if	voter 1 prefers A to B (i.e., $A\ R_1\ B$ and **not** $A\ I_1\ B$) (P for prefers),

same notations with A, C, or B, C, and with index 2 for voter 2.

Whatever the procedure is for extracting out of the preference schedules of 1 and 2 a resulting ranking for the group, where we denote again

$A\ R\ B$ (noindex)	if	the group outcome for A is higher than or equal to that for B;
$A\ I\ B$	if	A and B tie;
$A\ P\ B$	if	A is preferred to B,

we do want that procedure to satisfy the following requirements:

1. Suppose that the outcome for x and y (where x and y can be any two choices from A, B, C) is xRy, starting from preference schedules R_1, R_2. If we change the individual schedules so that x either moves up or stays the same in each, and y stays ranked at the same place, then the new social choice should still rank x higher or equal to y. The same statement is true for xPy: raising x in individual schedules preserves the higher ranking of x.

2. If the only change in the individual preference schedules concerns the ranking of the third alternative, but the individual relative rankings of x and y do not change, then the social ranking of x and y should not change.

3. For each choice of x, y among A, B, C, there exists some individual preference schedules that lead to xPy.

4. For each choice of x, y among A, B, C, there is some preference schedule in which xP_1y yet it does not follow that xPy (otherwise, 1 would be a dictator); the same holds for 2.

These are all very reasonable assumptions. But then, Arrow showed that:

> There exists **no** possible procedure for deriving R, P, I from the $R_1, P_1, I_1, R_2, P_2, I_2$ that satisfies all these assumptions.

In principle, you could do a computer search among all possibilities, and show that there is something wrong. But you can also argue the case as follows. We are going to derive a few consequences of our conditions $1 \to 4$ that will contradict each other.

- First of all, if the individual preference schedules are such that AP_1B and AP_2B, then we must have APB. (Same for other choices of the two alternatives to be ranked.)

  We know, from 3, that there are P_1', P_2' schedules for which the result is $AP'B$. In these two schedules, push alternative C to the bottom; this does not affect the relative rankings of A and B so the group ranking of A and B is unchanged. Now raise A to the top on both individual schedules if it isn't there yet. This does not affect the group preference of A (by condition 1). So we end up with new preference schedules P_1", P_2" for which AP_1"B, AP_2"B, and AP"B. Now we can pull up C in each of the two preference schedules to the place where it sits in P_1 and P_2. We can do this because AP_1B and AP_2B. Since the group ranking of A, B is only affected by the individual rankings of A and B (independently of where the individuals place C), and since we had AP"B, we must also have APB.

- Next, suppose that whenever AP_1B, BP_2A, we always find that APB. Then the following argument shows that this would mean that AP_1B implies APB, regardless of the schedule of 2 (which would be in contradiction with condition 4!).

  Take a preference schedule for 1 such that $AP_1'B$; we don't know what 2 does. If $BP_2'A$, then we already see that $AP'B$; so suppose $AP_2'B$ or AI_2B. If $AP_2'B$, then we also know that $AP'B$. It remains to look at $AI_2'B$. Now go over to different schedules R_1", R_2", P_1", P_2" for 1 and 2 by not changing anything to 1, but pushing A to the bottom in 2. Then AP_1"B and BP_2"A, so AP"B. To go back to the schedules with a single prime, we do nothing but raising A, so this does not affect the ranking of A over B (by 1). So $AP'B$ follows again.

So this contradiction with the dictator's rule means that it cannot be that AP_1B, BP_2A always implies APB (same for other pairs).

- Next, we show that AP_1B and BP_2A **must** imply AIB.

  Suppose that we have some schedules R_1', R_2', P_1', P_2' such that $AP_1'B$, $BP_2'A$, and $AP'B$. Since only relative rankings count, it would follow for **all** schedules that have AP_1B and BP_2A that APB. But that (as we just saw) is **not** allowed. So we can't have $AP'B$. Similarly, $BP'A$ is excluded. So it follows that $AI'B$ (same again for other pairs).

- Let us now look at the special preference schedule where

  $$\begin{array}{cc} AP_1B & BP_1C \\ CP_2A & AP_2B \end{array}$$

  (i.e., 1 ranks them A, B, C; 2 ranks them C, A, B). Then $AP_1B, AP_2B \Rightarrow APB$. But also

  $$\left.\begin{array}{l} AP_1C, CP_2A \Rightarrow AIC \\ BP_1C, CP_2B \Rightarrow BIC \end{array}\right\} \Rightarrow AIB\ .$$

  But A cannot at the same time outrank B and be equivalent to $B \Rightarrow$ contradiction!

We looked only at this special case (two voters, three alternatives), but the same argument holds for N voters, K alternatives, under similar conditions on the would-be "ideal" voting scheme.

(You don't even have to assume that all possible permutations of the K alternatives are "admissible" preferences for all voters. For instance, if you asked rankings of 20 or so menus by 100 people, then you would expect that the vegetarians among the group never rank meat dishes near the top, thus excluding some individual preference schedules. As long as there is one group of three alternatives where people can rank in any order, the argument will hold.)

For this surprising and unexpected result, Kenneth Arrow was awarded the Nobel Prize in Economics. You can read more about it in his book, *Social Choice and Individual Values* (on reserve in Fine Library) or in Steven Brams' book, *Rational Politics* (also on reserve in Fine Library—this account is more readable, but doesn't give the proof).

Arrow's theorem shows that there is no ideal voting system, that is, no system that will translate **all** possible individual preference schemes into a group preference scheme that is fair (in the sense of Arrow's conditions).

But maybe only a few of all possible schemes lead to a problem? When there is a Condorcet winner, it is certainly true that the first two conditions are satisfied:

- if, in all individual schemes, y stays put and x moves up or stays put, then the relative group ordering of x versus y cannot deteriorate.
- whether x or y wins in a pairwise comparison for the group, depends only on the relative positions of x and y in the individual rankings.

So trouble can arise only when there is no Condorcet winner. How often can this happen?

Let's look at the case of three voters, three alternatives.

Each voter has six possibilities for his/her ranking:

$$\begin{array}{cccccc} A & B & C & B & A & C \\ B & C & A & A & C & B \\ C & A & B & C & B & A \end{array}$$

The total number of possibilities is $6 \times 6 \times 6 = 216$.

How many of all these combinations lead to a voting paradox, i.e., to the absence of a Condorcet winner?

Take any of the six possibilities for voter 1, and let's check what possibilities for 2 and 3 exclude a Condorcet winner.

Example: Voter 1 picks

$$\begin{array}{c} C \\ A \\ B \end{array}$$

If voter 2 picks C first, then C will win. In order to avoid a Condorcet winner, 2 must therefore have ranked A or B first. Let us explore these two possibilities:

1. Voter 2 picks B first. If this voter's second choice is A then we have

$$\begin{array}{cc} C & B \\ A & A \\ B & C \end{array}$$

Then the third voter must have ranked A first, otherwise there would be a Condorcet winner. That means that voter 3 has two possible rankings only:

$$\begin{array}{ccc} A & \text{or} & A \\ B & & C \\ C & & B \end{array}$$

But in both cases there is a Condorcet winner, namely A. Our assumption that voter 2 picked A in second place always leads to controdiction, so voter 2's second choice **must** have been C:

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

It now follows that for 3 we have no choice:

First pick	A	
Second pick	B	(if 3 picked C in 2nd place,
Third pick	C	then C would be Condorcet winner)

2. In the second possibility, voter 2 picks A first.

Then we have again two possibilities for the second choice of 2: B or C

First possibility

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

Second possibility

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

both A and C have already outranked B twice. A has outranked C once, C has outranked A once. If 3 ranks A ahead of C, A is the Condorcet winner. If 3 ranks C ahead of A, C is the Condorcet winner. In this case there is therefore always a Condorcet winner.

$\Rightarrow$ only first possibility remains.

Now 3 must rank B first (otherwise there is a Condorcet winner) and C second (otherwise we have $\begin{array}{ccc} C & A & B \\ A & B & A \\ B & C & C \end{array}$ $\Rightarrow$ Condorcet winner).

There is thus only one possibility for 3 if we know that there is no Condorcet winner.

So, for each of voter 1's six possibilities, there are two for voter 2 and then no more choices for voter 3, that can lead to the absence of a Condorcet winner; in total we have thus twelve cases.

Only twelve out of 216 possibilities give a paradox, or $\frac{12}{216} = \frac{2}{36} = \frac{1}{18} \simeq .056 \rightarrow$ only 5.6% of cases. Maybe the "voting paradoxes" are much ado about almost nothing then?

Let us look at what happens for more general cases, for higher numbers of voters and of alternatives among which to choose. Here is a table of the percentage of preference schedules that give rise to a paradox for general numbers of voters and of alternatives:

Number of Alternatives	Number of Individuals						
	3	5	7	9	11	...	Limit
3	0.056	0.069	0.075	0.078	0.080	...	0.080
4	0.111	0.139	0.150	0.156	0.160	...	0.176
5	0.160	0.200	0.215	0.230	0.251	...	0.251
6	0.202	0.255	0.258	0.284	0.294	...	0.315
7	0.239	0.299	0.305	0.342	0.343	...	0.369
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Limit	1.000	1.000	1.000	1.000	1.000	...	1.000

For three alternatives and an arbitrary number of voters, it is never worse than .080, which is less than 1/12. But more alternatives rapidly increase the likelihood of a paradox ...

This is one reason why Congress in the U.S. (as well as in many other countries) uses voting procedures that are binary (i.e. only two alternatives are pitted against each other at any time; the winning alternative then leads another choice, ...). But that can lead to other problems, where the final outcome is not the one that would have received a majority vote. The following excerpt (from *Rational Politics*, by Steven Brams; on reserve in Fine Library) gives an example:

> ...concerns federal aid for school construction, which was considered by the U.S. House of Representatives in 1956 in terms of the following three alternatives: (1) an original bill, O, for grants-in-aid for school construction; (2) the bill with the so-called Powell amendment, A, which provided that no federal money be spent in states with segregated public schools; and (3) no bill, N, the status quo. As Riker reconstructed the situation, there were basically three groups of voters with the following preference scales:

- The passers (mostly southern Democrats) favored school aid, but they were so repelled by the Powell amendment that they preferred no bill to the amended bill. Thus, their preference scale was (O,N,A).
- The amenders (mostly northern Democrats) were prointegration, but if the Powell amendment were not successful, would still prefer school aid to no action. Thus, their preference scale was (A,O,N).
- The defeaters (mostly Republicans) preferred the status quo over either version of the bill, with the prointegration bill apparently more palatable than the original bill. Thus, their preference scale was (N,A,O).

The voting procedure used in Congress (and in many other voting bodies) is a binary procedure called the *amendment procedure*, whereby the set of outcomes is divided into two subsets, and each subset in turn into two further subsets, as shown in the *voting tree* depicted on the next page (read from top to bottom). In the example at hand, the procedure pits the subset of outcomes $\{A, N\}$ against the subset $\{O, N\}$. If the amendment is adopted, the second vote is on the amended bill (A) versus no bill (N); if not, the second vote is on the original bill (O) versus N. Defeat can thus occur in two different ways under this procedure.

The first vote on the Powell amendment passed, but the second vote on the amended bill failed. The first roll call reveals that the northern Democrats needed (and got) the support of Republicans, who, realizing that a majority probably favored the original bill, may have voted for the amendment $\{A, N\}$ instead of $\{O, N\}$ only so as to detach southern Democrat support on the vote for final passage (A versus N), causing it to fail. (On the other hand, if the Republicans had voted against the amendment initially, the vote for final passage would have been O versus N, and O would have won because both northern and southern Democrats would have supported it.)

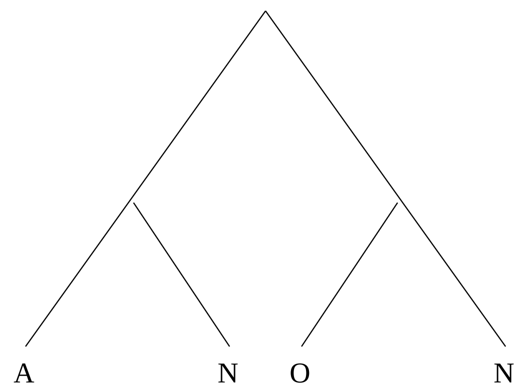
Since the Republicans did not introduce the Powell amendment, however, it seems unfair to blame them for using it to defeat the school-aid bill. ...

In fact, it was the northern Democrats, and not the Republicans, who had the opportunity to contrive an outcome that would have been preferred by them over N, but they failed to exploit this opportunity. Specifically, if they had voted for $\{O, N\}$ instead of their sincere choice $\{A, N\}$ on the amendment vote, $\{O, N\}$ would have been selected. Then outcome O would have been the choice of the voting body on the final vote, since two of the three groups of voters (southern Democrats and northern Democrats) would have preferred it to outcome N. ...

In 1957, a year after the Powell amendment had succeeded ($\{A, N\}$ defeated $\{O, N\}$) but final passage had failed (N defeated A), the bill in its original form got majority support in the House (O defeated N). This history is convincing evidence that a paradox of voting actually existed: O defeats N, N defeats A, and A defeats O.

(taken from p. 67-70 in *Rational Politics*, Steven Brams; this chapter contains another similar example)

Amendment Procedure:



first decision:
do we want to vote between {A,N} or {O,N}?

second decision:
vote on two alternatives in the subset chosen in first step.