Arrow's Theorem

Philosophy 444

4 November, 2019

In the other notes I was a bit informal about stating Arrow's Theorem, and this led to some confusion in class I think. So I thought I'd write down the formal version of it carefully here, though we won't be referring to it a lot.

Assume we are trying to construct a **social choice function** (scf) with the following features. There are n citizens and k options. Each citizen has a preference function that is a **total preorder** over the k options. The scf takes these n total preorders as inputs, and produces a total preorder representing the social preferences as output.

A total preorder is a function $\geq$ over the options satisfying:

Completeness For any x, y, either $x \geq y$ or $y \geq x$ (or both).
Transitivity For any x, y, z if $x \geq y$ and $y \geq z$ then $x \geq z$.

Intuitively, $x \geq y$ means that the relevant citizen (or group) thinks that x is at least as good as y. So we are assuming, in effect, that each citizen's preference ranking over the options is a total preorder.

For each citizen, there is a corresponding dictator function. The dictator function for citizen i is the scf that takes as inputs all the preference rankings, and outputs the rankings of citizen i, whatever the other $n - 1$ citizens say.

Arrow's Theorem is that these n dictator functions are the only functions that satisfy

Universal Domain For any possible input, the scf deterministically produces an output.
Independence of Irrelevant Alternatives (IIA) The social rank of x and y is a function of just how the n citizens rank x and y, not of how they rank these two options compared to other options.
Pareto If all citzens prefer x to y, then so does the social ranking

Universal domain does not rule out any kind of randomness. Since a total preorder allows for ties, when both $x \geq y$ and $y \geq x$, it could be that the scf outputs a ranking where x and y are tied at the top. And it could be that we use some kind of random process to choose between them in that circumstance. But it doesn't allow for any output where you would naturally use weighted probabilities - e.g., choose x with probability 0.8 and y with probability 0.2.

IIA is the really strong condition, the one that is violated by pretty much all formal systems that we use.

In the first version of his impossibility theorem, Arrow imposed a stronger condition than Pareto, namely monotonicity

Monotonicity Improving the position of x in one citizen's rankings, and making no other changes, cannot reduce x's position in the scf.

This condition is satisfied by plurality voting, but it is violated by runoff voting. (And, as a consequence, by instant runoff voting.) To see this, imagine a contest with three candidates, Left, Center and Right. (This example is similar to, but simpler than, the example of monotonicity violation on the wikipedia page about monotonicity.) There are 100 voters, and they are divided up as follows.

1st Choice	2nd Choice	Voters
Left	Center	45
Center	Left	12
Center	Right	15
Right	Center	28

Left gets 45, Center gets 27 and Right gets 28. So in the runoff it's Left vs Right, and Left wins 57-43.

Now imagine that two citizens are radicalised by a dramatic experience. They go from being Right-wing voters to Left-wing voters. The table now looks like

1st Choice	2nd Choice	Voters
Left	Center	47
Center	Left	12
Center	Right	15
Right	Center	26

Now Left gets 47, Center gets 27 and Right gets 26. So in the runoff it's Left vs Center, and Center wins 53-47.

Something seems off here, though it's not immediately obvious where the bad thing happened. But having these two voters start voting for Left doesn't seem like it should cause Left to lose.

Some critics of runoff voting (and IRV) think this is a serious reason to abandon it. I am a little suspicious of how widespread an issue this is. IRV is used for hundreds (if not thousands) of elections in Australia each year, and there are a handful of cases in all of electoral history where this might have been an issue. So I suspect it is more of a theoretical problem than a practical problem, but it is a surprising theoretical problem.