The Last Fair Vote

Why mathematics quietly broke democracy—and why that might be okay

The summer I turned seventeen, I voted for the first time.

It was a mock election at school—three candidates, a ballot box, the whole theatre. I ranked my preferences, dropped my slip into the box, and felt what I can only describe as a small, warm sense of civic virtue. The process felt clean. Mathematical, even. We count the votes; the winner wins. Simple.

What nobody told me—what I only discovered months later, hunched over a Wikipedia article at 11pm—is that a mathematician had proven, decades before I was born, that no voting system can ever be truly fair. Not 'fair' in a fuzzy, political sense. Fair in a precise, mathematical sense. Provably, necessarily, permanently impossible.

His name was Kenneth Arrow, and in 1951 he published a short, devastating theorem that earned him the Nobel Prize in Economics. I want to take you on the journey I took that evening, from naïve civic optimism to mathematical vertigo—and then, hopefully, to something resembling peace.

What Would a Fair Vote Even Look Like?

Before smashing something, you have to know what you're trying to build.

Imagine a society choosing between three options: a new park (P), a car park (C), or leaving the land empty (E). Everyone submits their personal ranking of the three—their *preference ordering*. A voting system is just a rule that takes all those individual rankings and produces a single collective ranking. Simple enough.

Now, what would we *want* such a system to do? Arrow identified three properties that seem, on reflection, completely non-negotiable.

The first is **unanimity**. If every single voter prefers P to C—every last one of them—then the collective ranking had better put P above C. If a society is unanimous and the vote ignores them, something has gone very wrong. This is sometimes called the Pareto condition, after the Italian economist Vilfredo Pareto.

The second is **independence of irrelevant alternatives** (IIA, if you want to sound impressive at parties). This says that the collective ranking of P versus C should depend *only* on how individuals rank P against C—not on where they've placed E. If we swap E around in everyone's personal lists without changing any P-vs-C opinions, the collective verdict on P versus C shouldn't budge. Otherwise, a candidate who can't win is somehow affecting who does win, which seems obviously wrong.

The third is **non-dictatorship**. There should be no single voter whose personal ranking automatically becomes the collective ranking, regardless of what everyone else thinks. A democracy run by one person's preferences is not, in fact, a democracy.

These three conditions are not radical demands. They are the bare minimum. They are the floor.

Arrow's theorem says you cannot have all three at once.

The Impossibility

I find the best way to feel the theorem—to really *feel* it rather than just accept it—is through a thought experiment about power.

Suppose we have any voting system that satisfies unanimity and IIA. We're going to show it must have a dictator.

Imagine voters standing in a long queue, numbered 1 through n. Initially, everyone prefers E to P (the park option, which we'll track throughout). By unanimity, the collective ranking also puts E above P. Now we're going to move P to the top of each voter's personal list, one at a time, starting from voter 1.

At some point, as we work through the queue, the collective ranking flips: E was above P, and now P is above E. There must be a specific voter—call them voter k—at whose desk the flip happened. Before we adjusted voter k's list, the collective said E $>$ P. After we adjusted it, the collective said P $>$ E. Voter k is pivotal.

Now here is the clever part. Through a careful series of further adjustments—exploiting IIA to isolate pairs of candidates—you can show that voter k's preferences determine the collective ranking for *every* pair of options, not just P versus E. Voter k is not just pivotal for one question. Voter k is a dictator.

The argument requires a little more machinery to run in full—it takes about two pages of careful logic—but the intuition is clear: if you want unanimity and IIA, you are forced to concentrate power into a single voter. You cannot spread it out. You cannot share it fairly.

The Condorcet Trap

Arrow's theorem is abstract. Let me show you the same impossibility wearing a different face.

The Marquis de Condorcet, an eighteenth-century French mathematician, noticed something alarming about majority voting. Consider three voters with these preferences:

- Voter 1: $P \succ C \succ E$
- Voter 2: $C \succ E \succ P$
- Voter 3: $E \succ P \succ C$

Now run majority votes between each pair. P beats C (voters 1 and 3 prefer P). C beats E (voters 1 and 2 prefer C). So surely P beats E? No: voters 2 and 3 prefer E to P.

We have P beats C, C beats E, *and* E beats P. A cycle. A loop with no winner. Majority preference is not transitive in the way that ordinary rankings are, and this is not a quirk of bad data—it is a fundamental feature of how individual preferences aggregate.

This is why IIA matters so much. If you introduce E into a race between P and C, you can—depending on your voting system—change who wins between P and C, even if no one has changed their mind about P versus C. The irrelevant alternative is not irrelevant at all. It is a crowbar.

So Is Democracy Broken?

I remember sitting back from my laptop that night feeling slightly cheated. Mathematics, I felt, had no business poking its nose into the ballot box.

But I've come to think the feeling was misplaced.

Arrow's theorem does not say that voting is pointless. It says that no voting system can simultaneously satisfy all three of those conditions. That is not a bug; it is a map. It tells us exactly where the compromises are hidden.

In practice, different systems choose different trade-offs. First-past-the-post abandons the ranking of multiple options entirely. Ranked-choice voting (instant runoff) gives up strict IIA. Approval voting lets people express support for multiple candidates but loses fine-grained preference information. Each sacrifice is real, and Arrow's theorem means no sacrifice-free option exists. The theorem does not pick a winner—it just makes sure we know what we're giving up when we choose.

There is something almost liberating in that. The quest for the perfect voting system was always chasing a mathematical ghost. The real question—which has always been a human question, not a mathematical one—is: *which* imperfection can we live with?

The Lesson in the Theorem

Mathematics has a peculiar habit of discovering things that feel like they belong to philosophy. Gödel showed that any sufficiently powerful logical system contains truths it cannot prove from within itself. Turing showed that no algorithm can generally determine whether another algorithm will ever stop running. And Arrow showed that the aggregation of individual preferences into collective ones must always, somewhere, contain a crack.

What strikes me about all three results is that they do not say the enterprise is hopeless. They say the enterprise is *bounded*—that there are walls we will never get past, and knowing where the walls are is itself a kind of knowledge.

When I voted in that school election—naively, warmly, with my small slip of paper—I was participating in a system that Arrow would later prove imperfect. But knowing a system is imperfect is very different from knowing it is useless. Bridges flex under load; that does not make them dangerous. Proofs have axioms; that does not make them arbitrary.

Democracy is a technology for turning disagreement into decisions. Like all technologies, it has trade-offs baked in. Arrow's gift was making those trade-offs legible.

The last fair vote was never real. But the search for something better very much is.

Further reading: Arrow's original paper is 'A Difficulty in the Concept of Social Welfare' (1950). For an accessible treatment, Amartya Sen's work on social choice theory is beautifully written. The Stanford Encyclopedia of Philosophy entry on 'Arrow's Theorem' is free, rigorous, and excellent.