Arrow

Arrow's Impossibility Theorem

What is Arrow's Impossibility Theorem?

Arrow's Impossibility Theorem (1951) demonstrates that no voting system can simultaneously satisfy a set of seemingly reasonable conditions for democratic decision-making.

The Central Question: Can we design a "perfect" democratic system that aggregates individual preferences into collective choices?

Arrow's Answer: No such system exists.

Individual Preference Rankings

Definition: An individual preference ranking is a complete ordering of all available alternatives from most preferred to least preferred.

Key Properties: - **Complete**: Every pair of alternatives can be compared - **Transitive**: If A > B and B > C, then A > C - **Antisymmetric**: If A > B, then not B > A

Example: Three voters choosing between candidates A, B, C - Voter 1: A > B > C - Voter 2: B > C > A

- Voter 3: C > A > B

The Challenge: How do we combine individual rankings into a single social ranking?

What We Want: A systematic method (social choice function) that takes individual preference profiles and produces a group preference ranking.

Examples of Attempted Solutions: - Majority rule - Plurality voting - Borda count - Condorcet method

The Problem: Arrow shows that no method can satisfy all our reasonable requirements simultaneously.

Arrow's Conditions for Democratic Choice

Arrow identified four conditions that seem essential for any fair democratic system:

1. **Unrestricted Domain** (U)
2. **Weak Pareto Principle** (P)
3. **Independence of Irrelevant Alternatives** (I)
4. **Non-dictatorship** (D)

We'll examine each condition, with special focus on Independence of Irrelevant Alternatives.

Condition 1: Unrestricted Domain (U)

What it means: The voting system must work for ANY possible combination of individual preference rankings.

Why it seems reasonable: We shouldn't restrict what preferences people are allowed to have in a democracy.

Example: The system must handle: - Unanimous preferences (everyone agrees) - Completely opposed preferences
- Cyclical preferences (A > B > C > A in majority comparisons) - Any other logically possible preference profile

Condition 2: Weak Pareto Principle (P)

What it means: If everyone prefers alternative X to alternative Y, then the group ranking should also prefer X to Y.

Why it seems reasonable: If there's unanimous agreement, the social choice should reflect that agreement.

Example: - If all voters rank A > B - Then the group ranking must have A > B - This seems like a minimal requirement for democracy

Condition 3: Independence of Irrelevant Alternatives (I)

What it means: The social ranking between any two alternatives should depend ONLY on how individuals rank those two alternatives, not on how they rank other alternatives.

Key Insight: Information about "irrelevant" third alternatives shouldn't change the relative ranking of two alternatives.

Why it seems reasonable: If we're deciding between A and B, it shouldn't matter what people think about C.

Independence of Irrelevant Alternatives: Examples

Example 1 - Violating IIA: - Initial preferences: Voter 1: A > B > C, Voter 2: B > A > C - Suppose social choice gives: A > B - Now C is removed from consideration - Preferences become: Voter 1: A > B, Voter 2: B > A
- IIA requires the social choice to still prefer A > B - Any system that now prefers B > A violates IIA

Example 2 - Why this matters: - In presidential elections, the presence of a third-party candidate can change who wins between the two main candidates - This violates IIA and can lead to strategic voting

Independence of Irrelevant Alternatives: Formal Statement

Formal Definition: For any two preference profiles R and R', if every individual has the same preference between alternatives x and y in both profiles, then the social choice between x and y must be the same in both profiles.

In other words: - Take any two alternatives: x and y - Consider two different voting scenarios - If each voter's preference between x and y is identical in both scenarios - Then the group's preference between x and y must also be identical

This rules out: Systems where the presence/absence of other alternatives affects pairwise comparisons.

Condition 4: Non-dictatorship (D)

What it means: No single individual's preferences should always determine the group ranking, regardless of what others prefer.

Why it seems reasonable: This is basic to democratic ideals - no one person should have absolute power.

Formal statement: There should be no individual i such that for every preference profile, if individual i prefers x to y, then the social choice prefers x to y.

Note: This doesn't prevent someone from being influential, just from being decisive in every case.

The Impossibility Theorem

Arrow's Theorem: There is no social choice function that simultaneously satisfies: - Unrestricted Domain (U) - Weak Pareto Principle (P) - Independence of Irrelevant Alternatives (I)
- Non-dictatorship (D)

What this means: Any democratic voting system must violate at least one of these seemingly reasonable conditions.

The Proof Strategy: Arrow showed that assuming all four conditions leads to a logical contradiction.

Why This Matters for Philosophy

Implications for Democratic Theory: - Challenges the idea of "the will of the people" - Shows that aggregating preferences is fundamentally problematic - Raises questions about the legitimacy of democratic outcomes

Broader Philosophical Questions: - What does this tell us about collective rationality? - Are there alternative approaches to social choice? - How should we respond to this impossibility?

Connection to Other Areas: - Philosophy of mind (collective vs. individual rationality) - Political philosophy (legitimacy of democratic institutions) - Ethics (how to make fair collective decisions)

Common Responses and Escape Routes

1. **Restrict the Domain**: Limit the types of preferences allowed (violates U)

2. **Accept Dictatorship**: In limited contexts, expert decision-making (violates D)

3. **Drop IIA**: Accept that context matters in social choice (violates I)

4. **Probabilistic Methods**: Use randomization or cardinal utilities

5. **Alternative Frameworks**: Deliberative democracy, constitutional constraints

The Debate Continues: Each response involves trade-offs and philosophical commitments.

Discussion Questions

1. Which of Arrow's four conditions would you be most willing to give up? Why?
2. Does Arrow's Theorem show that democracy is impossible, or just that perfect democracy is impossible?
3. How might this theorem apply to other collective decision-making contexts (committees, families, etc.)?
4. Do you think the Independence of Irrelevant Alternatives condition is actually reasonable? Can you think of cases where context should matter?
5. What does this theorem tell us about the nature of collective rationality?