Voting Impossibility Theorems

Aarush Aggarwal

Euler Circle IRPW

July 2025

How Should We Vote?

Key Questions:

- How do we aggregate individual preferences into a group decision?
- What makes a voting system "fair"?
- Can we design a system that avoids paradoxes and inconsistencies?

How Should We Vote?

Key Questions:

- How do we aggregate individual preferences into a group decision?
- What makes a voting system "fair"?
- Can we design a system that avoids paradoxes and inconsistencies?

Central Problem:

Can we design a voting rule that is both **fair** and **consistent** across all possible voter profiles?

The Need for Formal Voting Rules

What is a voting system?

- A **voting rule** maps a profile of individual preferences to a collective outcome.
- A **preference profile** consists of each voter's ranking of alternatives.

The Need for Formal Voting Rules

What is a voting system?

- A **voting rule** maps a profile of individual preferences to a collective outcome.
- A **preference profile** consists of each voter's ranking of alternatives.

Why is it hard?

- Voters may disagree or contradict each other.
- Simple majority comparisons can lead to paradoxes.

The Need for Formal Voting Rules

What is a voting system?

- A **voting rule** maps a profile of individual preferences to a collective outcome.
- A **preference profile** consists of each voter's ranking of alternatives.

Why is it hard?

- Voters may disagree or contradict each other.
- Simple majority comparisons can lead to paradoxes.

Goal: Define a system that transforms individual rationality into collective rationality.

Condorcet's Paradox: The Cycle Problem

Definition: An alternative x is a **Condorcet winner** if it beats every other alternative in pairwise majority votes.

Condorcet's Paradox: The Cycle Problem

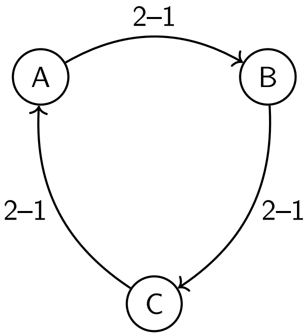
Definition: An alternative x is a **Condorcet winner** if it beats every other alternative in pairwise majority votes.
Condorcet Paradox: There may be no Condorcet winner! Pairwise majority outcomes can form a cycle, violating transitivity.

Condorcet's Paradox: The Cycle Problem

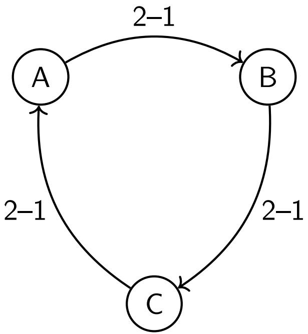
Definition: An alternative x is a **Condorcet winner** if it beats every other alternative in pairwise majority votes.
Condorcet Paradox: There may be no Condorcet winner! Pairwise majority outcomes can form a cycle, violating transitivity.
Example: 3 Voters, 3 Candidates

Voter	1st	2nd	3rd
p_1	A	B	C
p_2	B	C	A
p_3	C	A	B

Condorcet Majority Cycle

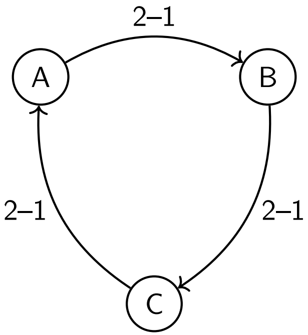


Condorcet Majority Cycle



Key Insight: Majority rule fails transitivity: A beats B, B beats C, but C beats A.

Condorcet Majority Cycle



Key Insight: Majority rule fails transitivity: A beats B, B beats C, but C beats A.

This **cycle** is the core of Condorcet's paradox.

Condorcet Paradox: Formal Definitions

Definition: Condorcet Winner
$x \in A$ is a Condorcet winner if:

$$\forall y \neq x, \quad |\{i : x \succ_i y\}| > |\{i : y \succ_i x\}|$$

Condorcet Paradox: Formal Definitions

Definition: Condorcet Winner

$x \in A$ is a Condorcet winner if:

$$\forall y \neq x, \quad |\{i : x \succ_i y\}| > |\{i : y \succ_i x\}|$$

Definition: Condorcet Cycle

There exists a sequence $x_1, x_2, \dots, x_k$ such that:

$$x_1 \succ_{\text{majority}} x_2 \succ_{\text{majority}} \cdots \succ x_k \succ x_1$$

Condorcet Paradox: Formal Definitions

Definition: Condorcet Winner
$x \in A$ is a Condorcet winner if:

$$\forall y \neq x, \quad |\{i : x \succ_i y\}| > |\{i : y \succ_i x\}|$$

Definition: Condorcet Cycle
There exists a sequence $x_1, x_2, \ldots, x_k$ such that:

$$x_1 \succ_{\text{majority}} x_2 \succ_{\text{majority}} \cdots \succ x_k \succ x_1$$

Theorem: A Condorcet cycle exists if and only if there is no Condorcet winner.

From Condorcet's Paradox to Arrow's Theorem

Condorcet's Paradox: Even majority rule can lead to cycles $\Rightarrow$ collective preferences become intransitive.

From Condorcet's Paradox to Arrow's Theorem

Condorcet's Paradox: Even majority rule can lead to cycles $\Rightarrow$ collective preferences become intransitive.

Key Insight: Simple, intuitive voting systems fail to guarantee consistent group decisions.

From Condorcet's Paradox to Arrow's Theorem

Condorcet's Paradox: Even majority rule can lead to cycles $\Rightarrow$ collective preferences become intransitive.

Key Insight: Simple, intuitive voting systems fail to guarantee consistent group decisions.

But is this just a fluke? Or something deeper?

From Condorcet's Paradox to Arrow's Theorem

Condorcet's Paradox: Even majority rule can lead to cycles ⇒ collective preferences become intransitive.

Key Insight: Simple, intuitive voting systems fail to guarantee consistent group decisions.

But is this just a fluke? Or something deeper?

Arrow's Contribution

Arrow showed that **no voting rule** can fully satisfy even a few basic fairness criteria — not just majority rule. *The problem is universal.*

Why Arrow's Theorem?

Goal: Find a social welfare function that fairly aggregates individual rankings into a collective ranking.

Why Arrow's Theorem?

Goal: Find a social welfare function that fairly aggregates individual rankings into a collective ranking.
Question: Can we find a rule that satisfies basic fairness conditions?

Why Arrow's Theorem?

Goal: Find a social welfare function that fairly aggregates individual rankings into a collective ranking.
Question: Can we find a rule that satisfies basic fairness conditions?
Arrow (1951) proved that this is **impossible** — under mild assumptions, any such rule must be dictatorial.

Why Arrow's Theorem?

Goal: Find a social welfare function that fairly aggregates individual rankings into a collective ranking.
Question: Can we find a rule that satisfies basic fairness conditions?
Arrow (1951) proved that this is **impossible** — under mild assumptions, any such rule must be dictatorial.
Let's first define the key axioms.

Axiom 1: Unrestricted Domain (U)

Definition: The social welfare function F must accept **every possible preference profile** as input.

Axiom 1: Unrestricted Domain (U)

Definition: The social welfare function F must accept **every possible preference profile** as input.
That is, for any tuple $(\succ_1, \ldots, \succ_n) \in L(A)^n$, the function $F()$ must be defined.

Axiom 1: Unrestricted Domain (U)

Definition: The social welfare function F must accept **every possible preference profile** as input.
That is, for any tuple $(\succ_1, \ldots, \succ_n) \in L(A)^n$, the function $F()$ must be defined.
Implication: The rule should work even when voters strongly disagree or form cycles.

Axiom 2: Pareto Efficiency (P)

Definition: If all voters prefer a to b, then society must also prefer a to b.

Axiom 2: Pareto Efficiency (P)

Definition: If all voters prefer a to b, then society must also prefer a to b.

$$\forall i \in N, \quad a \succ_i b \quad \Rightarrow \quad a \succ_F b$$

Axiom 2: Pareto Efficiency (P)

Definition: If all voters prefer a to b, then society must also prefer a to b.

$$\forall i \in N, \quad a \succ_i b \quad \Rightarrow \quad a \succ_F b$$

Motivation: If everyone agrees, the group must reflect that consensus.

Axiom 3: Independence of Irrelevant Alternatives (IIA)

Definition: The group's preference between a and b should depend **only** on how individuals rank a and b.

Axiom 3: Independence of Irrelevant Alternatives (IIA)

Definition: The group's preference between a and b should depend **only** on how individuals rank a and b.

$$\text{If } \succ_i \text{ and } \succ_i' \text{ agree on } a \text{ vs. } b, \Rightarrow F(), F(') \text{ agree on } a \text{ vs. } b$$

Axiom 3: Independence of Irrelevant Alternatives (IIA)

Definition: The group's preference between a and b should depend **only** on how individuals rank a and b.

$$\text{If } \succ_i \text{ and } \succ_i' \text{ agree on } a \text{ vs. } b, \Rightarrow F(), F(') \text{ agree on } a \text{ vs. } b$$

Motivation: Irrelevant candidates shouldn't change the outcome between a and b.

Axiom 4: Non-Dictatorship (ND)

Definition: No single voter i should have the power to determine the group ranking regardless of others.

Axiom 4: Non-Dictatorship (ND)

Definition: No single voter i should have the power to determine the group ranking regardless of others.

$$\text{There is no } i \in N \text{ such that } \forall,\ F() = \succ_i$$

Axiom 4: Non-Dictatorship (ND)

Definition: No single voter i should have the power to determine the group ranking regardless of others.

$$\text{There is no } i \in N \text{ such that } \forall,\ F() = \succ_i$$

Motivation: A fair rule must not give total power to one individual.

Summary of Arrow's Axioms

- **U — Unrestricted Domain:** All profiles allowed.
- **P — Pareto Efficiency:** Consensus is respected.
- **IIA — Independence of Irrelevant Alternatives:** Only relevant preferences matter.
- **ND — Non-Dictatorship:** No one person decides everything.

Summary of Arrow's Axioms

- **U — Unrestricted Domain:** All profiles allowed.
- **P — Pareto Efficiency:** Consensus is respected.
- **IIA — Independence of Irrelevant Alternatives:** Only relevant preferences matter.
- **ND — Non-Dictatorship:** No one person decides everything.

Can we satisfy all of them together?

Arrow's Impossibility Theorem (Formal Statement)

Theorem (Arrow, 1951):
Let $|A| \geq 3$, $|N| \geq 2$. Then, no social welfare function $F : L(A)^n \to L(A)$ satisfies:

- Unrestricted Domain (U)
- Pareto Efficiency (P)
- Independence of Irrelevant Alternatives (IIA)
- Non-Dictatorship (ND)

Arrow's Impossibility Theorem (Formal Statement)

Theorem (Arrow, 1951):
Let $|A| \geq 3$, $|N| \geq 2$. Then, no social welfare function $F : L(A)^n \to L(A)$ satisfies:

- Unrestricted Domain (U)
- Pareto Efficiency (P)
- Independence of Irrelevant Alternatives (IIA)
- Non-Dictatorship (ND)

Any rule satisfying the first three must be a dictatorship.

Proof Sketch: Arrow's Impossibility Theorem

Goal: Show that any voting rule satisfying U, P, and IIA must be a dictatorship.

Proof Sketch: Arrow's Impossibility Theorem

Goal: Show that any voting rule satisfying U, P, and IIA must be a dictatorship.

Idea: Use Decisive Coalitions.

1. **Start with the full group** — by Pareto, all voters together are decisive.

Proof Sketch: Arrow's Impossibility Theorem

Goal: Show that any voting rule satisfying U, P, and IIA must be a dictatorship.

Idea: Use Decisive Coalitions.

1. **Start with the full group** — by Pareto, all voters together are decisive.
2. **Use IIA + transitivity** to show this decisiveness "spreads" to new pairs (Field Expansion).

Proof Sketch: Arrow's Impossibility Theorem

Goal: Show that any voting rule satisfying U, P, and IIA must be a dictatorship.

Idea: Use Decisive Coalitions.

1. **Start with the full group** — by Pareto, all voters together are decisive.
2. **Use IIA + transitivity** to show this decisiveness "spreads" to new pairs (Field Expansion).
3. **Then shrink the group** step-by-step using the Contraction Lemma.

Proof Sketch: Arrow's Impossibility Theorem

Goal: Show that any voting rule satisfying U, P, and IIA must be a dictatorship.

Idea: Use Decisive Coalitions.

1. **Start with the full group** — by Pareto, all voters together are decisive.
2. **Use IIA + transitivity** to show this decisiveness "spreads" to new pairs (Field Expansion).
3. **Then shrink the group** step-by-step using the Contraction Lemma.
4. Eventually you reach a **single decisive voter** — a **dictator**.

Formal Proof of Arrow's Theorem: Decisive Coalitions

Definition: A coalition $S \subseteq N$ is *decisive* for (x, y) if for any profile where all $i \in S$ have $x \succ_i y$, we get $x \succ_F y$ regardless of preferences outside S.

Lemma 1 (Field Expansion): If S is decisive for (x, y), then S is decisive for (x, z) and (z, y) for any $z \in A$.

Formal Proof of Arrow's Theorem: Decisive Coalitions

Definition: A coalition $S \subseteq N$ is *decisive* for (x, y) if for any profile where all $i \in S$ have $x \succ_i y$, we get $x \succ_F y$ regardless of preferences outside S.

Lemma 1 (Field Expansion): If S is decisive for (x, y), then S is decisive for (x, z) and (z, y) for any $z \in A$.
Idea: Construct profiles leveraging Pareto, IIA, and transitivity to expand the decisiveness.

Formal Proof of Arrow's Theorem: Decisive Coalitions

Definition: A coalition $S \subseteq N$ is *decisive* for (x, y) if for any profile where all $i \in S$ have $x \succ_i y$, we get $x \succ_F y$ regardless of preferences outside S.

Lemma 1 (Field Expansion): If S is decisive for (x, y), then S is decisive for (x, z) and (z, y) for any $z \in A$.
Idea: Construct profiles leveraging Pareto, IIA, and transitivity to expand the decisiveness.
Lemma 2 (Contraction): If S is globally decisive and $|S| \geq 2$, then $\exists T \subsetneq S$ such that T is also globally decisive.

Formal Proof of Arrow's Theorem: Decisive Coalitions

Definition: A coalition $S \subseteq N$ is *decisive* for (x, y) if for any profile where all $i \in S$ have $x \succ_i y$, we get $x \succ_F y$ regardless of preferences outside S.

Lemma 1 (Field Expansion): If S is decisive for (x, y), then S is decisive for (x, z) and (z, y) for any $z \in A$.
Idea: Construct profiles leveraging Pareto, IIA, and transitivity to expand the decisiveness.
Lemma 2 (Contraction): If S is globally decisive and $|S| \geq 2$, then $\exists T \subsetneq S$ such that T is also globally decisive.
Consequence: By recursively contracting N, we reach a singleton voter i^* who is decisive for all pairs — a dictator.

Arrow's Theorem: Final Formal Statement

Theorem (Arrow, 1951):
Let $|A| \geq 3$ and $|N| \geq 2$. Then, no social welfare function $F : L(A)^n \rightarrow L(A)$ satisfies:

- Unrestricted Domain (U)
- Pareto Efficiency (P)
- Independence of Irrelevant Alternatives (IIA)
- Non-Dictatorship (ND)

Arrow's Theorem: Final Formal Statement

Theorem (Arrow, 1951):
Let $|A| \geq 3$ and $|N| \geq 2$. Then, no social welfare function $F : L(A)^n \to L(A)$ satisfies:

- Unrestricted Domain (U)
- Pareto Efficiency (P)
- Independence of Irrelevant Alternatives (IIA)
- Non-Dictatorship (ND)

Proof Sketch:

1. Use Pareto to show N is globally decisive.
2. Apply contraction to reduce to a singleton decisive voter.
3. Singleton decisiveness $\Rightarrow$ dictatorship.
4. Contradicts ND.

The Gibbard–Satterthwaite Theorem

Theorem (Gibbard, Satterthwaite):
Let $|A| \geq 3$, and let $f : L(A)^n \to A$ be a deterministic, onto, and strategy-proof social choice function.

Then f must be a **dictatorship**.

The Gibbard–Satterthwaite Theorem

Theorem (Gibbard, Satterthwaite):
Let $|A| \geq 3$, and let $f : L(A)^n \to A$ be a deterministic, onto, and strategy-proof social choice function.

Then f must be a **dictatorship**.
Implication: Even if we drop IIA and ask only for truthful voting (strategy-proofness), the outcome must still be dictated by a single voter.

Idea Behind the Proof

Key Concepts:

- **Strategy-proofness:** No voter can benefit by misreporting.
- **Pivotal voter:** A single voter can change the outcome between a and b by modifying their ranking.
- **Monotonicity:** Raising a winning candidate in a ballot shouldn't cause them to lose.

Outline:

1. Start with a profile where $f() = a$.
2. Gradually raise b in one voter's ranking until $f(') = b$.
3. First voter to cause the switch is **pivotal**.
4. Repeat for all pairs (a, b), (a, c), etc.
5. That voter determines all pairwise outcomes $\Rightarrow$ dictator.

Interpretation of Gibbard–Satterthwaite

What does this mean for democracy?

- Any attempt to ensure voters report truthfully — under deterministic and complete systems — ends up violating fairness.
- The theorem complements Arrow's: it proves strategy-proofness is incompatible with fairness too.
- Dictatorship is the only solution immune to manipulation in this setting.

Interpretation of Gibbard–Satterthwaite

What does this mean for democracy?

- Any attempt to ensure voters report truthfully — under deterministic and complete systems — ends up violating fairness.
- The theorem complements Arrow's: it proves strategy-proofness is incompatible with fairness too.
- Dictatorship is the only solution immune to manipulation in this setting.

Conclusion: Strategic resistance $\wedge$ full domain $\wedge$ determinism $\Rightarrow$ Dictatorship

What is Metric Distortion?

Definition

Metric Distortion is a concept used to evaluate how much a voting outcome deviates from the optimal outcome, based on a set of voter preferences. It is the ratio of the distance between the selected outcome and the voters' preferences to the distance between the ideal outcome and the voters' preferences.

- It quantifies the difference between the ideal candidate (according to the voters' preferences) and the winner selected by the voting rule.
- The idea is to **measure fairness** in voting systems: how close the chosen outcome is to the best possible one.

Key Terms: Preference Profile and Voting Rule

- **Preference Profile**: A collection of individual rankings from all voters. Each voter ranks the candidates in order of preference.

$$\mathcal{P} = \begin{cases} \text{Voter 1: } A > B > C \\ \text{Voter 2: } B > C > A \\ \text{Voter 3: } C > A > B \end{cases}$$

- **Voting Rule**: A *voting rule* is a function f that maps a preference profile $\mathcal{P}$ to a collective outcome, i.e., $f : \mathcal{P} \to \mathcal{O}$. Examples include Plurality Rule, Borda Count, and Condorcet Method.

- **Metric Space**: A set of candidates $A = \{a_1, a_2, \ldots, a_m\}$, where a **distance function** $d(v, c)$ measures how far a candidate c is from a voter v.

The Formula for Metric Distortion

Mathematical Definition: Given a preference profile $\mathcal{P}$, a set of

candidates A, and a voting rule f, the **distortion** of the voting rule is defined as:

$$\text{Distortion}(f) = \frac{d(\mathcal{P}, f(\mathcal{P}))}{d(\mathcal{P}, \mathcal{P}^*)}$$

Explanation

- The **ideal social cost** $D(\mathcal{P}, \mathcal{P}^*)$ is the distance between the voters' preferences and the best possible outcome (ideal outcome).
- The **actual social cost** $D(\mathcal{P}, f(\mathcal{P}))$ is the distance between the voters' preferences and the outcome that the voting rule selects.

Calculating Actual and Ideal Social Cost

Formula: Ideal Social Cost

The **ideal social cost** $D(\mathcal{P}, \mathcal{P}^*)$ is the total distance between the voters' preferences $\mathcal{P}$ and the ideal outcome $\mathcal{P}^*$. For each voter p_i, the ideal candidate C_i^* is the one closest to them.

$$D(\mathcal{P}, \mathcal{P}^*) = \sum_{i=1}^{n} D(p_i, C_i^*)$$

Formula: Actual Social Cost

The **actual social cost** $D(\mathcal{P}, f(\mathcal{P}))$ is the total distance between the voters' preferences $\mathcal{P}$ and the outcome selected by the voting rule $f(\mathcal{P})$.

$$D(\mathcal{P}, f(\mathcal{P})) = \sum_{i=1}^{n} D(p_i, f(\mathcal{P}))$$

Geometric Space and Voter/Candidate Positions

Consider a **geometric space** in $\mathbb{R}^2$.

$\mathcal{P}$ contains all voter preferences where each $p_i = (x_i, y_i)$ is the position of voter i in the space. Similarly, each candidate C_j is positioned at (x_j, y_j) in $\mathbb{R}^2$.

Formula: Distance Function

The **distance function** between a voter $p_i = (x_i, y_i)$ and a candidate $C_j = (x_j, y_j)$ is given by the **Euclidean distance**:

$$D(p_i, C_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

Geometric Plane with Voters and Candidates

Here's a visual of the Gemoteric Space