Arrow's Impossibility Theorem

Aaron Roth

University of Pennsylvania

Fall 2026

Overview

- Motivation: In RLHF, diverse human comparisons become a single training objective.

Overview

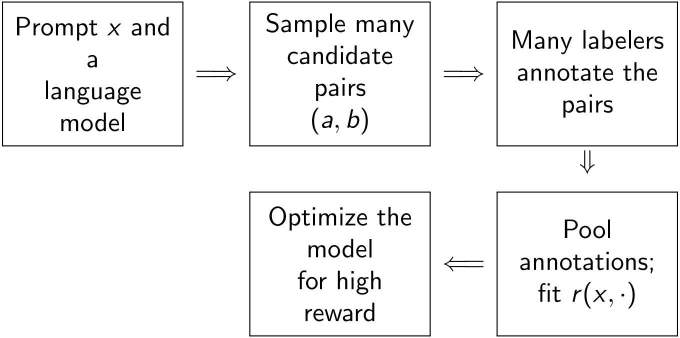
- Motivation: In RLHF, diverse human comparisons become a single training objective.
- Two requirements we might ask of an ideal aggregation rule.

Overview

- Motivation: In RLHF, diverse human comparisons become a single training objective.
- Two requirements we might ask of an ideal aggregation rule.
- Arrow's theorem: these requirements force one person's ranking to determine the outcome.

Human Comparisons Become a Training Objective

Reinforcement learning from human feedback (RLHF) uses this pipeline:



Reward Fitting Contains an Aggregation Problem

Fix one prompt x and a finite set C of candidate completions.

Reward Fitting Contains an Aggregation Problem

Fix one prompt x and a finite set C of candidate completions.

- Many labelers compare pairs $a, b \in C$.
- A scalar reward $r(x, \cdot)$ assigns one number to each completion.
- It therefore induces a *single ranking* of all candidates in C.

Reward Fitting Contains an Aggregation Problem

Fix one prompt x and a finite set C of candidate completions.

- Many labelers compare pairs $a, b \in C$.
- A scalar reward $r(x, \cdot)$ assigns one number to each completion.
- It therefore induces a *single ranking* of all candidates in C.

Whose ranking should it be?

A Running Example

One math prompt produces three responses:

Type	Primary concern	Ranking
1	Concision	$a \succ b \succ c$
2	Formal rigor	$b \succ c \succ a$
3	Intuition	$c \succ a \succ b$

A Running Example

One math prompt produces three responses:

Type	Primary concern	Ranking
1	Concision	$a \succ b \succ c$
2	Formal rigor	$b \succ c \succ a$
3	Intuition	$c \succ a \succ b$

- Every type has a complete transitive ordering over completions.
- The three types occur in equal proportions.

A Running Example

One math prompt produces three responses:

Type	Primary concern	Ranking
1	Concision	$a \succ b \succ c$
2	Formal rigor	$b \succ c \succ a$
3	Intuition	$c \succ a \succ b$

- Every type has a complete transitive ordering over completions.
- The three types occur in equal proportions.

How might we combine these rankings into one?

Pairwise Majority Decides Each Contest Separately

Definition

The *pairwise majority relation* ranks a above b, written $a \succ_M b$, when more voters rank a above b than rank b above a.

Pairwise Majority Decides Each Contest Separately

Definition

The *pairwise majority relation* ranks a above b, written $a \succ_M b$, when more voters rank a above b than rank b above a.

It uses the most direct possible rule for each pair:

> compare the number favoring a with the number favoring b.

Pairwise Majority Decides Each Contest Separately

Definition

The *pairwise majority relation* ranks a above b, written $a \succ_M b$, when more voters rank a above b than rank b above a.

It uses the most direct possible rule for each pair:

compare the number favoring a with the number favoring b.

But do the winners of the separate pairwise contests fit together into a single transitive ordering?

Coherent Individuals Can Produce a Cycle

Type	Ranking
1	$a \succ b \succ c$
2	$b \succ c \succ a$
3	$c \succ a \succ b$

Coherent Individuals Can Produce a Cycle

Type	Ranking
1	$a \succ b \succ c$
2	$b \succ c \succ a$
3	$c \succ a \succ b$

1. Types 1 and 3 prefer a to b: $a \succ_M b$.

Coherent Individuals Can Produce a Cycle

Type	Ranking
1	$a \succ b \succ c$
2	$b \succ c \succ a$
3	$c \succ a \succ b$

1. Types 1 and 3 prefer a to b: $a \succ_M b$.
2. Types 1 and 2 prefer b to c: $b \succ_M c$.

Coherent Individuals Can Produce a Cycle

Type	Ranking
1	$a \succ b \succ c$
2	$b \succ c \succ a$
3	$c \succ a \succ b$

1. Types 1 and 3 prefer a to b: $a \succ_M b$.
2. Types 1 and 2 prefer b to c: $b \succ_M c$.
3. Types 2 and 3 prefer c to a: $c \succ_M a$.

Coherent Individuals Can Produce a Cycle

Type	Ranking
1	$a \succ b \succ c$
2	$b \succ c \succ a$
3	$c \succ a \succ b$

1. Types 1 and 3 prefer a to b: $a \succ_M b$.
2. Types 1 and 2 prefer b to c: $b \succ_M c$.
3. Types 2 and 3 prefer c to a: $c \succ_M a$.

Thus pairwise majority requires

$$a \succ_M b \succ_M c \succ_M a.$$

Coherent Individuals Can Produce a Cycle

Type	Ranking
1	$a \succ b \succ c$
2	$b \succ c \succ a$
3	$c \succ a \succ b$

1. Types 1 and 3 prefer a to b: $a \succ_M b$.
2. Types 1 and 2 prefer b to c: $b \succ_M c$.
3. Types 2 and 3 prefer c to a: $c \succ_M a$.

Thus pairwise majority requires

$$a \succ_M b \succ_M c \succ_M a.$$

No scalar reward can represent all three comparisons:

$$r(a) > r(b) > r(c) > r(a).$$

Why Ask for a Full Ranking?

- In an election, we may only need to select one winner.

Why Ask for a Full Ranking?

- In an election, we may only need to select one winner.
- A scalar reward model does more: it compares *every pair* of completions.

Why Ask for a Full Ranking?

- In an election, we may only need to select one winner.
- A scalar reward model does more: it compares *every pair* of completions.
- If ties are broken consistently, it ranks every pair without cycles.

Why Ask for a Full Ranking?

- In an election, we may only need to select one winner.
- A scalar reward model does more: it compares *every pair* of completions.
- If ties are broken consistently, it ranks every pair without cycles.
- Arrow asks whether we can construct such an order from individual preferences.

Why Ask for a Full Ranking?

- In an election, we may only need to select one winner.
- A scalar reward model does more: it compares *every pair* of completions.
- If ties are broken consistently, it ranks every pair without cycles.
- Arrow asks whether we can construct such an order from individual preferences.

Matches the object learned by scalar RLHF better than it matches the goal of elections, its classical application.

The Aggregation Object

Let C be a finite set of alternatives.

The Aggregation Object

Let C be a finite set of alternatives.

Definition

A *strict preference order* $\succ$ on C is *complete*—it ranks every pair of distinct alternatives—and *transitive*. Write $\mathcal{L}(C)$ for the set of all such orders.

The Aggregation Object

Let C be a finite set of alternatives.

Definition

A *strict preference order* $\succ$ on C is *complete*—it ranks every pair of distinct alternatives—and *transitive*. Write $\mathcal{L}(C)$ for the set of all such orders.

The input is a preference profile

$$\succ = (\succ_1, \ldots, \succ_n) \in \mathcal{L}(C)^n.$$

The Aggregation Object

Let C be a finite set of alternatives.

Definition

A *strict preference order* $\succ$ on C is *complete*—it ranks every pair of distinct alternatives—and *transitive*. Write $\mathcal{L}(C)$ for the set of all such orders.

The input is a preference profile

$$\boldsymbol{\succ} = (\succ_1, \ldots, \succ_n) \in \mathcal{L}(C)^n.$$

Definition

An *Arrovian social welfare function*, or *social ranking rule*, is

$$F : \mathcal{L}(C)^n \longrightarrow \mathcal{L}(C).$$

Write $\succ_F^{\boldsymbol{\succ}}$ for its output at profile $\boldsymbol{\succ}$.

The Aggregation Object

Let C be a finite set of alternatives.

Definition

A *strict preference order* $\succ$ on C is *complete*—it ranks every pair of distinct alternatives—and *transitive*. Write $\mathcal{L}(C)$ for the set of all such orders.

The input is a preference profile

$$\boldsymbol{\succ} = (\succ_1, \ldots, \succ_n) \in \mathcal{L}(C)^n.$$

Definition

An *Arrovian social welfare function*, or *social ranking rule*, is

$$F : \mathcal{L}(C)^n \longrightarrow \mathcal{L}(C).$$

Write $\succ_F^{\boldsymbol{\succ}}$ for its output at profile $\boldsymbol{\succ}$.

Axiom 1: Pareto Unanimity

Definition

F satisfies *Pareto unanimity* if for every profile and every pair $a, b \in C$,

$$a \succ_i b \text{ for every voter } i \quad \Longrightarrow \quad a \succ_F b.$$

Axiom 1: Pareto Unanimity

Definition

F satisfies *Pareto unanimity* if for every profile and every pair $a, b \in C$,

$$a \succ_i b \text{ for every voter } i \quad \Longrightarrow \quad a \succ_F b.$$

If everyone agrees on one comparison, society should not reverse it.

Axiom 1: Pareto Unanimity

Definition

F satisfies *Pareto unanimity* if for every profile and every pair $a, b \in C$,

$$a \succ_i b \text{ for every voter } i \quad \Longrightarrow \quad a \succ_F b.$$

If everyone agrees on one comparison, society should not reverse it.

It does not specify what to do when voters disagree.

Axiom 2: Independence of Irrelevant Alternatives

Definition

F satisfies *IIA* if the social comparison between a and b depends only on the voters' comparisons between a and b.
Formally, if profiles $\succ$ and $\succ'$ satisfy

$$a \succ_i b \iff a \succ'_i b \qquad \text{for every } i,$$

then

$$a \succ_F^{\succ} b \iff a \succ_F^{\succ'} b.$$

Axiom 2: Independence of Irrelevant Alternatives

Definition

F satisfies *IIA* if the social comparison between a and b depends only on the voters' comparisons between a and b.
Formally, if profiles $\succ$ and $\succ'$ satisfy

$$a \succ_i b \iff a \succ'_i b \qquad \text{for every } i,$$

then

$$a \succ_F^{\succ} b \iff a \succ_F^{\succ'} b.$$

Voters may move other alternatives anywhere—only their a–b comparisons must remain fixed.

What We Want to Exclude: Dictatorship

Definition

F is a *dictatorship* if there is a voter d such that, at every profile and for every pair a, b,

$$a \succ_d b \implies a \succ_F b.$$

What We Want to Exclude: Dictatorship

Definition

F is a *dictatorship* if there is a voter d such that, at every profile and for every pair a, b,

$$a \succ_d b \quad \Longrightarrow \quad a \succ_F b.$$

Because both orders are complete, the social ranking exactly reproduces voter d's ranking.

What We Want to Exclude: Dictatorship

Definition

F is a *dictatorship* if there is a voter d such that, at every profile and for every pair a, b,

$$a \succ_d b \implies a \succ_F b.$$

Because both orders are complete, the social ranking exactly reproduces voter d's ranking.

A dictatorship satisfies Pareto and IIA—but ignores everyone except d.

Borda Count

With m alternatives, each voter awards

$$m-1, m-2, \ldots, 0$$

points from first place to last place.

Borda Count

With m alternatives, each voter awards

$$m-1, m-2, \ldots, 0$$

points from first place to last place.

Definition

The *Borda count* sums these scores across voters and ranks alternatives by their totals, with a fixed rule for breaking ties.

Borda Count

With m alternatives, each voter awards

$$m-1, m-2, \dots, 0$$

points from first place to last place.

Definition

The *Borda count* sums these scores across voters and ranks alternatives by their totals, with a fixed rule for breaking ties.

Equivalently: draw a voter uniformly and draw a distinct opponent uniformly. An alternative's Borda score is proportional to its probability of winning that comparison.

Borda Count

With m alternatives, each voter awards

$$m - 1, m - 2, \ldots, 0$$

points from first place to last place.

Definition

The *Borda count* sums these scores across voters and ranks alternatives by their totals, with a fixed rule for breaking ties.

Equivalently: draw a voter uniformly and draw a distinct opponent uniformly. An alternative's Borda score is proportional to its probability of winning that comparison.

Borda always produces a transitive score ranking—but other alternatives can affect the social comparison between a and b.

Borda Violates IIA

	Voter 1	Voter 2	Voter 3	Voter 4	Scores (a, b, c)
P	$a > c > b$	$a > c > b$	$b > a > c$	$b > a > c$	$(6, 4, 2)$

Borda Violates IIA

	Voter 1	Voter 2	Voter 3	Voter 4	Scores (a, b, c)
P	$a > c > b$	$a > c > b$	$b > a > c$	$b > a > c$	$(6, 4, 2)$
P'	$a > b > c$	$a > b > c$	$b > c > a$	$b > c > a$	$(4, 6, 2)$

Borda Violates IIA

	Voter 1	Voter 2	Voter 3	Voter 4	Scores (a, b, c)
P	$a > c > b$	$a > c > b$	$b > a > c$	$b > a > c$	$(6, 4, 2)$
P'	$a > b > c$	$a > b > c$	$b > c > a$	$b > c > a$	$(4, 6, 2)$

- Every voter has the same a–b comparison at P and P'.
- Borda ranks a above b at P, but b above a at P'.
- Moving only the third alternative changes the social a–b comparison.

The Natural Rules Each Miss Something

Rule	Strict order?	Pareto?	IIA?
Dictatorship	✓	✓	✓

The Natural Rules Each Miss Something

Rule	Strict order?	Pareto?	IIA?
Dictatorship	✓	✓	✓
Pairwise majority	×	✓	✓

The Natural Rules Each Miss Something

Rule	Strict order?	Pareto?	IIA?
Dictatorship	✓	✓	✓
Pairwise majority	×	✓	✓
Borda count	✓	✓	×

The Natural Rules Each Miss Something

Rule	Strict order?	Pareto?	IIA?
Dictatorship	✓	✓	✓
Pairwise majority	×	✓	✓
Borda count	✓	✓	×

- We saw pairwise majority cycle.
- We saw Borda reverse a and b after only c moved.
- Is there a fourth rule with all three checkmarks that is not a dictatorship?

Arrow's Impossibility Theorem

Theorem (Arrow)

Let $|C| \geq 3$. *Every social ranking rule*

$$F : \mathcal{L}(C)^n \longrightarrow \mathcal{L}(C)$$

defined on the unrestricted domain of strict preference profiles and satisfying Pareto unanimity and IIA is a dictatorship.

Proof Roadmap

Fix one alternative b.

1. If every voter puts b at an extreme, society must put b at an extreme.

Proof Roadmap

Fix one alternative b.

1. If every voter puts b at an extreme, society must put b at an extreme.
2. Move b from bottom to top one voter at a time; some voter's move changes the social ranking.

Proof Roadmap

Fix one alternative b.

1. If every voter puts b at an extreme, society must put b at an extreme.
2. Move b from bottom to top one voter at a time; some voter's move changes the social ranking.
3. This only identifies a pivotal voter at one pair of profiles.

Proof Roadmap

Fix one alternative b.

1. If every voter puts b at an extreme, society must put b at an extreme.
2. Move b from bottom to top one voter at a time; some voter's move changes the social ranking.
3. This only identifies a pivotal voter at one pair of profiles.
4. Use IIA to show that the same voter dictates every pair not involving b at every profile.

Proof Roadmap

Fix one alternative b.

1. If every voter puts b at an extreme, society must put b at an extreme.
2. Move b from bottom to top one voter at a time; some voter's move changes the social ranking.
3. This only identifies a pivotal voter at one pair of profiles.
4. Use IIA to show that the same voter dictates every pair not involving b at every profile.
5. Use a third alternative to extend control to pairs involving b.

Step 1: An Extreme Alternative Stays Extreme

Lemma

Suppose every voter places b either at the very top or at the very bottom of their ranking. Then the social order also places b at the top or at the bottom.

Step 1: An Extreme Alternative Stays Extreme

Lemma

Suppose every voter places b either at the very top or at the very bottom of their ranking. Then the social order also places b at the top or at the bottom.

Suppose not. Then for some alternatives a, c,

$$a \succ_F b \succ_F c.$$

Step 1: An Extreme Alternative Stays Extreme

Lemma

Suppose every voter places b either at the very top or at the very bottom of their ranking. Then the social order also places b at the top or at the bottom.

Suppose not. Then for some alternatives a, c,

$$a \succ_F b \succ_F c.$$

We will modify the profile so that:

1. every comparison involving b remains unchanged;
2. every voter ranks c above a.

Proof of the Extreme-Alternative Lemma

Only the restriction to $\{a, b, c\}$ is shown:

$$\begin{array}{lccc} b \text{ at top} & b \succ_i a,\ b \succ_i c & \rightsquigarrow & b \succ'_i c \succ'_i a \\ b \text{ at bottom} & a \succ_i b,\ c \succ_i b & \rightsquigarrow & c \succ'_i a \succ'_i b \end{array}$$

Proof of the Extreme-Alternative Lemma

Only the restriction to $\{a, b, c\}$ is shown:

$$\begin{array}{lccc} b \text{ at top} & b \succ_i a,\ b \succ_i c & \rightsquigarrow & b \succ'_i c \succ'_i a \\ b \text{ at bottom} & a \succ_i b,\ c \succ_i b & \rightsquigarrow & c \succ'_i a \succ'_i b \end{array}$$

IIA preserves both social comparisons involving b:

$$a \succ'_F b \succ'_F c.$$

Proof of the Extreme-Alternative Lemma

Only the restriction to $\{a, b, c\}$ is shown:

b at top	$b \succ_i a,\ b \succ_i c$	$\rightsquigarrow$	$b \succ'_i c \succ'_i a$
b at bottom	$a \succ_i b,\ c \succ_i b$	$\rightsquigarrow$	$c \succ'_i a \succ'_i b$

IIA preserves both social comparisons involving b:

$$a \succ'_F b \succ'_F c.$$

Pareto gives $c \succ'_F a$.

Proof of the Extreme-Alternative Lemma

Only the restriction to $\{a, b, c\}$ is shown:

b at top	$b \succ_i a,\ b \succ_i c$	$\rightsquigarrow$	$b \succ'_i c \succ'_i a$
b at bottom	$a \succ_i b,\ c \succ_i b$	$\rightsquigarrow$	$c \succ'_i a \succ'_i b$

IIA preserves both social comparisons involving b:

$$a \succ'_F b \succ'_F c.$$

Pareto gives $c \succ'_F a$. But transitivity of $a \succ'_F b \succ'_F c$ gives $a \succ'_F c$—a contradiction.

Step 2: One Voter Makes b Switch Extremes

Fix every voter's ranking of $C \setminus \{b\}$.

Step 2: One Voter Makes b Switch Extremes

Fix every voter's ranking of $C \setminus \{b\}$.
For $k = 0, \ldots, n$, form profile $P^{(k)}$:

- voters $1, \ldots, k$ put b at the top;
- voters $k+1, \ldots, n$ put b at the bottom.

Step 2: One Voter Makes b Switch Extremes

Fix every voter's ranking of $C \setminus \{b\}$.
For $k = 0, \ldots, n$, form profile $P^{(k)}$:

- voters $1, \ldots, k$ put b at the top;
- voters $k+1, \ldots, n$ put b at the bottom.

$$P^{(0)} \longrightarrow P^{(1)} \longrightarrow \cdots \longrightarrow P^{(n)}.$$

Step 2: One Voter Makes b Switch Extremes

Fix every voter's ranking of $C \setminus \{b\}$.
For $k = 0, \ldots, n$, form profile $P^{(k)}$:

- voters $1, \ldots, k$ put b at the top;
- voters $k+1, \ldots, n$ put b at the bottom.

$$P^{(0)} \longrightarrow P^{(1)} \longrightarrow \cdots \longrightarrow P^{(n)}.$$

Pareto puts b socially at the bottom at $P^{(0)}$ and at the top at $P^{(n)}$.

Step 2: One Voter Makes b Switch Extremes

Fix every voter's ranking of $C \setminus \{b\}$.
For $k = 0, \ldots, n$, form profile $P^{(k)}$:

- voters $1, \ldots, k$ put b at the top;
- voters $k+1, \ldots, n$ put b at the bottom.

$$P^{(0)} \longrightarrow P^{(1)} \longrightarrow \cdots \longrightarrow P^{(n)}.$$

Pareto puts b socially at the bottom at $P^{(0)}$ and at the top at $P^{(n)}$. The extreme-alternative lemma says b is always socially top or bottom.

Step 2: One Voter Makes b Switch Extremes

Fix every voter's ranking of $C \setminus \{b\}$.
For $k = 0, \ldots, n$, form profile $P^{(k)}$:

- voters $1, \ldots, k$ put b at the top;
- voters $k+1, \ldots, n$ put b at the bottom.

$$P^{(0)} \longrightarrow P^{(1)} \longrightarrow \cdots \longrightarrow P^{(n)}.$$

Pareto puts b socially at the bottom at $P^{(0)}$ and at the top at $P^{(n)}$. The extreme-alternative lemma says b is always socially top or bottom. So for some first index d, changing only voter d makes b jump from social bottom to social top.

What Have We Proved About Voter d?

Let
$$P^- := P^{(d-1)}, \qquad P^+ := P^{(d)}.$$

Then:

	voter d places b	society places b
P^-	bottom	bottom
P^+	top	top

What Have We Proved About Voter d?

Let

$$P^- := P^{(d-1)}, \qquad P^+ := P^{(d)}.$$

Then:

	voter d places b	society places b
P^-	bottom	bottom
P^+	top	top

Only voter d changed.

What Have We Proved About Voter d?

Let

$$P^- := P^{(d-1)}, \qquad P^+ := P^{(d)}.$$

Then:

	voter d places b	society places b
P^-	bottom	bottom
P^+	top	top

Only voter d changed. We call d *extremely pivotal for* b at the pair (P^-, P^+).

What Have We Proved About Voter d?

Let

$$P^{-} := P^{(d-1)}, \qquad P^{+} := P^{(d)}.$$

Then:

	voter d places b	society places b
P^{-}	bottom	bottom
P^{+}	top	top

Only voter d changed. We call d *extremely pivotal for* b at the pair (P^{-}, P^{+}).

But this conclusion concerns only these two profiles.
Why must d control the social ranking at every other profile?

Step 3: Extend Control to Every Profile

Lemma
For any distinct alternatives $a, c \neq b$ *and any profile* R,

$$a \succ_d c \quad \Longrightarrow \quad a \succ_F^R c.$$

Step 3: Extend Control to Every Profile

Lemma

For any distinct alternatives $a, c \neq b$ and any profile R,

$$a \succ_d c \quad \Longrightarrow \quad a \succ_F^R c.$$

So the locally pivotal voter d dictates every comparison that does not involve b—at *every* profile.

Step 3: Extend Control to Every Profile

Lemma

For any distinct alternatives $a, c \neq b$ and any profile R,

$$a \succ_d c \quad \Longrightarrow \quad a \succ_F^R c.$$

So the locally pivotal voter d dictates every comparison that does not involve b—at *every* profile.

To transfer this control to an arbitrary profile, we will construct an auxiliary profile that matches each pivotal profile on one relevant pair.

Construct an Auxiliary Profile

Fix an arbitrary target profile R with $a \succ_d c$.

Construct an Auxiliary Profile

Fix an arbitrary target profile R with $a \succ_d c$.
Construct a profile Q:

1. Voters $i < d$ put b at the top.

Construct an Auxiliary Profile

Fix an arbitrary target profile R with $a \succ_d c$.
Construct a profile Q:

1. Voters $i < d$ put b at the top.
2. Voter d ranks $a \succ_d b \succ_d c$.

Construct an Auxiliary Profile

Fix an arbitrary target profile R with $a \succ_d c$.
Construct a profile Q:

1. Voters $i < d$ put b at the top.
2. Voter d ranks $a \succ_d b \succ_d c$.
3. Voters $i > d$ put b at the bottom.

Construct an Auxiliary Profile

Fix an arbitrary target profile R with $a \succ_d c$.
Construct a profile Q:

1. Voters $i < d$ put b at the top.
2. Voter d ranks $a \succ_d b \succ_d c$.
3. Voters $i > d$ put b at the bottom.
4. Every voter's a–c comparison in Q matches their comparison in R.

Match Q to P^- and P^+ One Pair at a Time

At Q, the individual a–b comparisons match those at P^-:

	$i < d$	$i \geq d$
a vs. b	$b \succ_i a$	$a \succ_i b$

Match Q to P^- and P^+ One Pair at a Time

At Q, the individual a–b comparisons match those at P^-:

	$i < d$	$i \geq d$
a vs. b	$b \succ_i a$	$a \succ_i b$

Since b is socially bottom at P^-, IIA gives

$$a \succ_F^Q b.$$

Match Q to P^- and P^+ One Pair at a Time

At Q, the individual a–b comparisons match those at P^-:

	$i < d$	$i \geq d$
a vs. b	$b \succ_i a$	$a \succ_i b$

Since b is socially bottom at P^-, IIA gives

$$a \succ_F^Q b.$$

At Q, the individual b–c comparisons match those at P^+:

	$i \leq d$	$i > d$
b vs. c	$b \succ_i c$	$c \succ_i b$

Match Q to P^- and P^+ One Pair at a Time

At Q, the individual a–b comparisons match those at P^-:

	$i < d$	$i \geq d$
a vs. b	$b \succ_i a$	$a \succ_i b$

Since b is socially bottom at P^-, IIA gives

$$a \succ_F^Q b.$$

At Q, the individual b–c comparisons match those at P^+:

	$i \leq d$	$i > d$
b vs. c	$b \succ_i c$	$c \succ_i b$

Since b is socially top at P^+, IIA gives

$$b \succ_F^Q c.$$

Finish the Local-to-Global Step

We have shown that at the auxiliary profile Q,

$$a \succ_F^Q b \succ_F^Q c.$$

Finish the Local-to-Global Step

We have shown that at the auxiliary profile Q,

$$a \succ_F^Q b \succ_F^Q c.$$

Transitivity gives

$$a \succ_F^Q c.$$

Finish the Local-to-Global Step

We have shown that at the auxiliary profile Q,

$$a \succ_F^Q b \succ_F^Q c.$$

Transitivity gives

$$a \succ_F^Q c.$$

By construction, every voter has the same a–c comparison in Q and in the arbitrary target profile R.

Finish the Local-to-Global Step

We have shown that at the auxiliary profile Q,

$$a \succ_F^Q b \succ_F^Q c.$$

Transitivity gives

$$a \succ_F^Q c.$$

By construction, every voter has the same a–c comparison in Q and in the arbitrary target profile R. IIA therefore gives

$$a \succ_F^R c.$$

Finish the Local-to-Global Step

We have shown that at the auxiliary profile Q,

$$a \succ_F^Q b \succ_F^Q c.$$

Transitivity gives

$$a \succ_F^Q c.$$

By construction, every voter has the same a–c comparison in Q and in the arbitrary target profile R. IIA therefore gives

$$a \succ_F^R c.$$

The same voter d dictates every pair away from b, at every profile.

Step 4: Comparisons Involving b

We still need to show that voter d dictates comparisons of the form a versus b.

Step 4: Comparisons Involving b

We still need to show that voter d dictates comparisons of the form a versus b.

Lemma

For every $a \neq b$ and every profile R,

$$a \succ_d b \iff a \succ_F^R b.$$

Step 4: Comparisons Involving b

We still need to show that voter d dictates comparisons of the form a versus b.

Lemma

For every $a \neq b$ and every profile R,

$$a \succ_d b \quad \Longleftrightarrow \quad a \succ_F^R b.$$

Choose a third alternative $c \notin \{a, b\}$.

Case 1: Voter d Ranks a Above b

Fix a target profile R with $a \succ_d b$.

Case 1: Voter d Ranks a Above b

Fix a target profile R with $a \succ_d b$.
Construct Q preserving every voter's a–b comparison from R, and:

- put voter d's local order at $a \succ_d c \succ_d b$;
- make every voter rank c above b.

Case 1: Voter d Ranks a Above b

Fix a target profile R with $a \succ_d b$.
Construct Q preserving every voter's a–b comparison from R, and:

- put voter d's local order at $a \succ_d c \succ_d b$;
- make every voter rank c above b.

The previous lemma applies to $a, c \neq b$:

$$a \succ_d c \quad \Longrightarrow \quad a \succ_F^Q c.$$

Case 1: Voter d Ranks a Above b

Fix a target profile R with $a \succ_d b$.
Construct Q preserving every voter's a–b comparison from R, and:

- put voter d's local order at $a \succ_d c \succ_d b$;
- make every voter rank c above b.

The previous lemma applies to $a, c \neq b$:

$$a \succ_d c \quad \Longrightarrow \quad a \succ_F^Q c.$$

Pareto gives $c \succ_F^Q b$, so transitivity gives $a \succ_F^Q b$.

Case 1: Voter d Ranks a Above b

Fix a target profile R with $a \succ_d b$.
Construct Q preserving every voter's a–b comparison from R, and:

- put voter d's local order at $a \succ_d c \succ_d b$;
- make every voter rank c above b.

The previous lemma applies to $a, c \neq b$:

$$a \succ_d c \quad \Longrightarrow \quad a \succ_F^Q c.$$

Pareto gives $c \succ_F^Q b$, so transitivity gives $a \succ_F^Q b$. IIA transfers the a–b comparison from Q to R.

Case 2: Voter d Ranks b Above a

Fix a target profile R with $b \succ_d a$.

Case 2: Voter d Ranks b Above a

Fix a target profile R with $b \succ_d a$.
Construct Q preserving every voter's a–b comparison from R, and:

- put voter d's local order at $b \succ_d c \succ_d a$;
- make every voter rank b above c.

Case 2: Voter d Ranks b Above a

Fix a target profile R with $b \succ_d a$.
Construct Q preserving every voter's a–b comparison from R, and:

- put voter d's local order at $b \succ_d c \succ_d a$;
- make every voter rank b above c.

Pareto gives $b \succ_F^Q c$.

Case 2: Voter d Ranks b Above a

Fix a target profile R with $b \succ_d a$.
Construct Q preserving every voter's a–b comparison from R, and:

- put voter d's local order at $b \succ_d c \succ_d a$;
- make every voter rank b above c.

Pareto gives $b \succ_F^Q c$. The previous lemma gives $c \succ_F^Q a$.

Case 2: Voter d Ranks b Above a

Fix a target profile R with $b \succ_d a$.
Construct Q preserving every voter's a–b comparison from R, and:

- put voter d's local order at $b \succ_d c \succ_d a$;
- make every voter rank b above c.

Pareto gives $b \succ_F^Q c$. The previous lemma gives $c \succ_F^Q a$.
Transitivity gives $b \succ_F^Q a$, and IIA transfers this comparison to R.

Conclusion of the Proof

Starting from one locally pivotal event, we proved:

1. voter d dictates every pair not involving b at every profile;

Conclusion of the Proof

Starting from one locally pivotal event, we proved:

1. voter d dictates every pair not involving b at every profile;
2. voter d also dictates every pair involving b at every profile.

Conclusion of the Proof

Starting from one locally pivotal event, we proved:

1. voter d dictates every pair not involving b at every profile;
2. voter d also dictates every pair involving b at every profile.

Therefore, for every profile R and every pair a, c,

$$a \succ_d c \quad \Longrightarrow \quad a \succ_F^R c.$$

Conclusion of the Proof

Starting from one locally pivotal event, we proved:

1. voter d dictates every pair not involving b at every profile;
2. voter d also dictates every pair involving b at every profile.

Therefore, for every profile R and every pair a, c,

$$a \succ_d c \quad \Longrightarrow \quad a \succ_F^R c.$$

The same fixed voter d controls every comparison at every profile. Thus F is a dictatorship. $\square$

How Can We Escape the Impossibility?

Arrow tells us which assumption must give way:

1. **Restrict preferences.** For example, place alternatives on a line and require each voter to rank them by distance from a favorite point; these preferences are *single-peaked*.

How Can We Escape the Impossibility?

Arrow tells us which assumption must give way:

1. **Restrict preferences.** For example, place alternatives on a line and require each voter to rank them by distance from a favorite point; these preferences are *single-peaked*.
2. **Permit non-transitive collective comparisons.** Pairwise majority may cycle.

How Can We Escape the Impossibility?

Arrow tells us which assumption must give way:

1. **Restrict preferences.** For example, place alternatives on a line and require each voter to rank them by distance from a favorite point; these preferences are *single-peaked*.
2. **Permit non-transitive collective comparisons.** Pairwise majority may cycle.
3. **Change the output.** Return a probability distribution over alternatives rather than a full ranking, as a *maximal lottery* does.

How Can We Escape the Impossibility?

Arrow tells us which assumption must give way:

1. **Restrict preferences.** For example, place alternatives on a line and require each voter to rank them by distance from a favorite point; these preferences are *single-peaked*.
2. **Permit non-transitive collective comparisons.** Pairwise majority may cycle.
3. **Change the output.** Return a probability distribution over alternatives rather than a full ranking, as a *maximal lottery* does.
4. **Use richer information.** Ask voters to rank lotteries; the von Neumann–Morgenstern axioms then yield numerical expected utilities.

How Can We Escape the Impossibility?

Arrow tells us which assumption must give way:

1. **Restrict preferences.** For example, place alternatives on a line and require each voter to rank them by distance from a favorite point; these preferences are *single-peaked*.
2. **Permit non-transitive collective comparisons.** Pairwise majority may cycle.
3. **Change the output.** Return a probability distribution over alternatives rather than a full ranking, as a *maximal lottery* does.
4. **Use richer information.** Ask voters to rank lotteries; the von Neumann–Morgenstern axioms then yield numerical expected utilities.
5. **Abandon IIA.** Borda and other score-based rules use the positions of other alternatives.

Thanks!

See you next class!