Optimized Democracy

Fall 2026 | Lecture 4

Distortion

Ariel Procaccia | Harvard University

MOTIVATION

- The goal of social choice is to aggregate individual preferences or opinions towards a socially desirable outcome
- Axioms attempt to capture social desirability, but they don't identify the "best" rule
- Perhaps we can quantify how socially desirable a rule is through social welfare?
- The challenge is that we don't know the voters' cardinal preferences — only their ordinal preferences

MOTIVATION

1	2	3	4
a	b	c	d
b	a	b	b
c	c	d	c
d	d	a	a

Preference profile

	1	2	3	4
a	1/4	0	0	0
b	1/4	1	1/2	1/2
c	1/4	0	1/2	0
d	1/4	0	0	1/2

Utility profile

- W.l.o.g. the plurality winner is a
- But supposed the preference profile is induced by the utility profile on the right
- Social welfare of a is $1/4$, whereas that of b is $9/4$ — 9 times as high!

UTILITARIAN DISTORTION

- As usual, we have a set of voters N of size n and a set of alternatives A of size m
- Each voter $i \in N$ has a utility function $u_i: A \rightarrow \mathbb{R}^+$
- $\boldsymbol{u} = (u_1, \dots, u_n)$ is a utility profile
- Assume that for all $i \in N$, $\sum_{x \in A} u_i(x) = 1$
- u_i induces a ranking σ_i, denoted $u_i \rhd \sigma_i$, if

$$x \succ_{\sigma_i} y \Rightarrow u_i(x) \geq u_i(y)$$

UTILITARIAN DISTORTION

- Denote the (utilitarian) social welfare of $x \in A$ by $\mathrm{sw}(x, \boldsymbol{u}) = \sum_{i \in N} u_i(x)$
- For a preference profile $\boldsymbol{\sigma}$ and $x \in A$, the utilitarian distortion of x at $\boldsymbol{\sigma}$ is
$$\mathrm{dist}_U(x, \boldsymbol{\sigma}) = \max_{y \in A} \max_{\boldsymbol{u} \rhd \boldsymbol{\sigma}} \frac{\mathrm{sw}(y, \boldsymbol{u})}{\mathrm{sw}(x, \boldsymbol{u})}$$
- The utilitarian distortion of $f: \mathcal{L}^n \to A$ is
$$\mathrm{dist}_U(f) = \max_{\boldsymbol{\sigma} \in \mathcal{L}^n} \mathrm{dist}_U(f(\boldsymbol{\sigma}), \boldsymbol{\sigma})$$

UTILITARIAN DISTORTION: EXAMPLE

1	2	3
a	b	b
c	a	c
b	c	a

Suppose we choose a
In the profile $\boldsymbol{\sigma}$

	1	2	3
a	1/3	0	0
b	1/3	1	1
c	1/3	0	0

$$\max_{\boldsymbol{u} \rhd \boldsymbol{\sigma}} \frac{\mathrm{sw}(b, \boldsymbol{u})}{\mathrm{sw}(a, \boldsymbol{u})} = 7$$

	1	2	3
a	1/3	0	0
b	1/3	1	1/2
c	1/3	0	1/2

$$\max_{\boldsymbol{u} \rhd \boldsymbol{\sigma}} \frac{\mathrm{sw}(c, \boldsymbol{u})}{\mathrm{sw}(a, \boldsymbol{u})} = \frac{5}{2}$$

$$\Rightarrow \mathrm{dist}_U(a, \boldsymbol{\sigma}) = 7$$

Poll 1

What is $\mathrm{dist}_U(b, \boldsymbol{\sigma})$?

- 1
- In [1,2)
- In [2,3)
- In $[3, \infty)$

LOWER BOUND

- Theorem: For any $f: \mathcal{L} \to A$, $\text{dist}_U(f) = \Omega(m^2)$
- Proof:
  - Let $\boldsymbol{\sigma}$ such that the voters are partitioned into sets $N_1, \ldots, N_{m-1}$, each of size (roughly) $n/(m-1)$
  - The voters in N_i rank a_i first and a_m second
  - It holds that $\text{dist}(a_m, \boldsymbol{\sigma}) = \infty$ — why?
  - If $f(\boldsymbol{\sigma}) = a_i \neq a_m$, consider $\boldsymbol{u}$ such that N_i have utility $1/m$ for all alternatives, and other voters have utility ½ for the top two choices
  - It holds that
    $$\text{sw}(a_i, \boldsymbol{u}) = \frac{n}{m-1} \cdot \frac{1}{m}, \quad \text{sw}(a_m, \boldsymbol{u}) \geq \frac{1}{2} \cdot \left(n - \frac{n}{m-1}\right) = \Omega(n)$$
  - Overall, it holds that
    $$\text{dist}_U(f) \geq \text{dist}_U(f(\boldsymbol{\sigma}), \boldsymbol{\sigma}) \geq \frac{\text{sw}(a_m, \boldsymbol{u})}{\text{sw}(a_i, \boldsymbol{u})} = \Omega(m^2) \blacksquare$$

UPPER BOUND

- Which voting rule might achieve a good — ideally $O(m^2)$ — upper bound on distortion?
- Let's try to rule out a few candidates

Poll 2

Which rule has unbounded distortion?

- Plurality
- Borda count
- Both rules
- Neither one

UPPER BOUND

- **Theorem:** $\mathrm{dist}_U(\text{plurality}) = O(m^2)$
- **Proof:**
  - Given a preference profile $\boldsymbol{\sigma}$, let the plurality winner be x
  - x is ranked first by at least n/m voters
  - Let $\boldsymbol{u} \rhd \boldsymbol{\sigma}$, then
    $$\mathrm{sw}(x, \boldsymbol{u}) \geq \frac{n}{m} \cdot \frac{1}{m} = \frac{n}{m^2}$$
  - For any $y \in A$, $\mathrm{sw}(y, \boldsymbol{u}) \leq n$
  - It follows that
    $$\mathrm{dist}_U(\text{plurality}) \leq \frac{n}{n/m^2} = m^2 \blacksquare$$

INSTANCE OPTIMALITY

- The instance-optimal rule $f^\star$ satisfies
  $$f^\star(\boldsymbol{\sigma}) \in \text{argmin}_{x \in A} \text{dist}_U(x, \boldsymbol{\sigma})$$
- It holds that $\text{dist}_u(f^\star) = \Theta(m^2)$
- This rule is easy to compute:

1	2	3
a	b	b
c	a	c
b	c	a

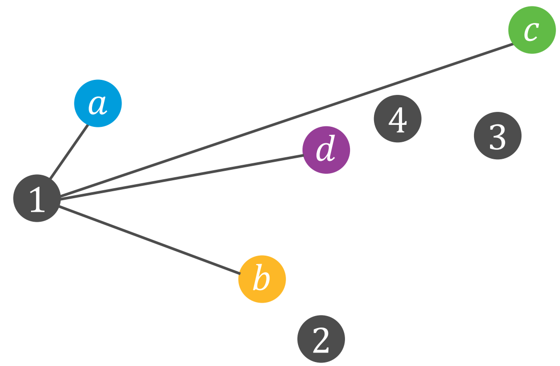
	1	2	3
a	1/3	0	0
b	1/3	1	1/2
c	1/3	0	1/2

Construct a utility profile that maximizes $\frac{\text{sw}(c,\boldsymbol{u})}{\text{sw}(a,\boldsymbol{u})}$

METRIC DISTORTION

1	2	3	4
a	b	c	d
b	d	d	c
d	a	b	b
c	c	a	a

Preference profile



Metric space

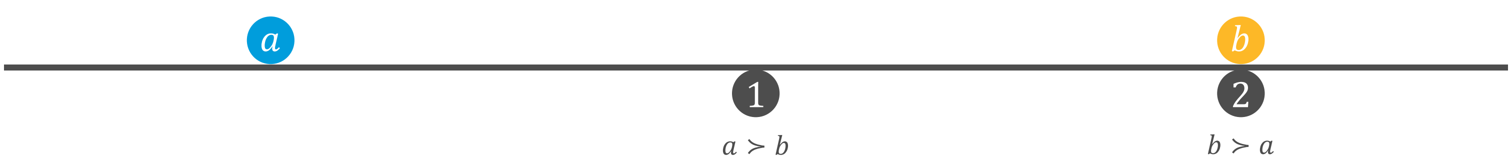
- Voters and alternatives lie in a latent metric space with metric ρ
- The preference profile is induced by the metric
- We are interested in minimizing the social cost, denoted $\mathrm{sc}(x,\rho) = \sum_{i\in N} \rho(i,x)$

METRIC DISTORTION

- Assume that $\boldsymbol{\sigma}$ is induced by a metric ρ satisfying:
  - $\forall x, y \in A, x \succ_{\sigma_i} y \Rightarrow \rho(i,x) \leq \rho(i,y)$
  - Symmetry: $\forall \alpha, \beta,\ \rho(\alpha,\beta) = \rho(\beta,\alpha)$
  - Triangle inequality: $\forall \alpha, \beta, \gamma,$
    $$\rho(\alpha,\beta) \leq \rho(\alpha,\gamma) + \rho(\gamma,\beta)$$
- Redefine distortion of x at $\boldsymbol{\sigma}$:
  $$\mathrm{dist}_M(x,\boldsymbol{\sigma}) = \max_{y\in A} \max_{\rho \rhd \boldsymbol{\sigma}} \frac{\mathrm{sc}(x,\rho)}{\mathrm{sc}(y,\rho)}$$
- As before, the distortion of f is
  $$\mathrm{dist}_M(f) = \max_{\boldsymbol{\sigma}\in\mathcal{L}^n} \mathrm{dist}_M(f(\boldsymbol{\sigma}),\boldsymbol{\sigma})$$

LOWER BOUND

- **Theorem:** For all $f: \mathcal{L}^n \to A$, $\text{dist}_M(f) \geq 3$
- **Proof:**
  - Consider a preference profile $\boldsymbol{\sigma}$ where $a \succ_{\sigma_1} b$ and $b \succ_{\sigma_2} a$
  - W.l.o.g. $f(\boldsymbol{\sigma}) = a$
  - Then consider the metric space below ■



UPPER BOUND

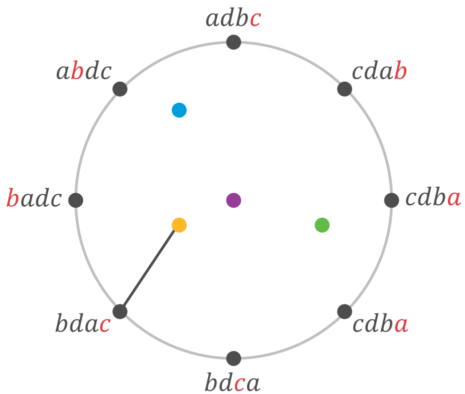
- The PluralityVeto rule works as follows:
  - The score of each alternative is initialized to its plurality score
  - One by one (in arbitrary order), voters decrement the score of their least preferred surviving alternative
  - Alternatives whose score is 0 are eliminated
  - Last alternative to be vetoed wins
- Theorem: $\text{dist}_M(\text{PluralityVeto}) \leq 3$
- Proof from "the Book"

PROOF OF THEOREM

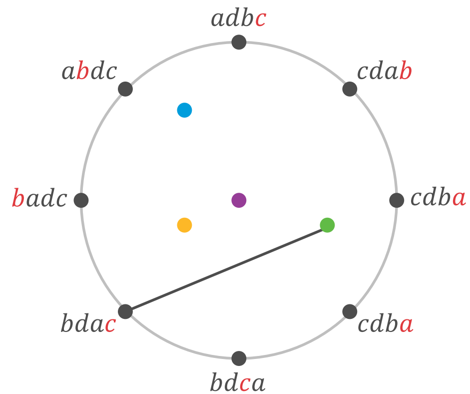
Let z_i be the alternative vetoed by $i \in N$, let $x^\star$ be the PluralityVeto winner, let N_x be the voters ranking x first, and let $y \in A$

$$
\begin{aligned}
\sum_{i \in N} \rho(i, x^\star) &\leq \sum_{i \in N} \rho(i, z_i) && (x^\star \succcurlyeq_{\sigma_i} z_i \text{ for all } i \in N) \\
&\leq \sum_{i \in N} \big(\rho(i, y) + \rho(y, z_i)\big) && (\text{triangle inequality}) \\
&= \sum_{i \in N} \rho(i, y) + \sum_{x \in A} \sum_{j \in N_x} \rho(y, x) && (\#\text{vetoes of } x \text{ is } |N_x|) \\
&\leq \sum_{i \in N} \rho(i, y) + \sum_{x \in A} \sum_{j \in N_x} \big(\rho(j, y) + \rho(j, x)\big) && (\text{triangle inequality}) \\
&\leq \sum_{i \in N} \rho(i, y) + \sum_{x \in A} \sum_{j \in N_x} 2\rho(j, y) && (i \in N_x \Rightarrow x \succcurlyeq_{\sigma_i} y) \\
&= 3 \sum_{i \in N} \rho(i, y) \quad \blacksquare
\end{aligned}
$$

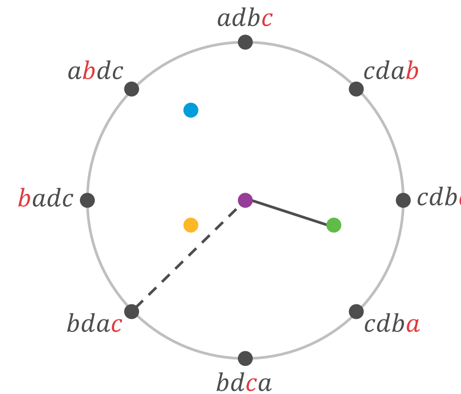
PROOF OF THEOREM



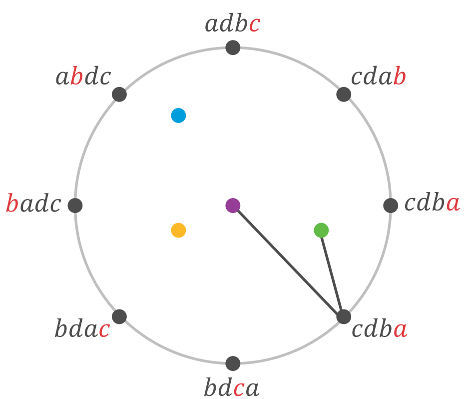
$\rho(i, x^\star)$



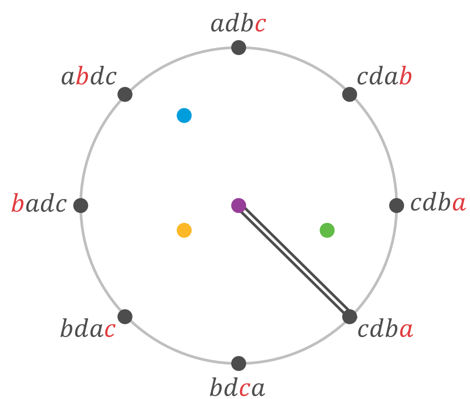
$\rho(i, z_i)$



$\rho(i, d) + \rho(d, z_i)$



$\rho(j, d) + \rho(j, c), j \in N_c$



$2\rho(j, d)$

$$
\begin{aligned}
\sum_{i\in N} \rho(i, x^\star) &\le \sum_{i\in N} \rho(i, z_i) \\
&\le \sum_{i\in N} \big(\rho(i, y) + \rho(y, z_i)\big) \\
&\le \sum_{i\in N} \rho(i, y) + \sum_{x\in A} \sum_{j\in N_x} \rho(y, x) \\
&\le \sum_{i\in N} \rho(i, y) + \sum_{x\in A} \sum_{j\in N_x} \big(\rho(j, y) + \rho(j, x)\big) \\
&\le \sum_{i\in N} \rho(i, y) + \sum_{x\in A} \sum_{j\in N_x} 2\rho(j, y) \\
&= 3 \sum_{i\in N} \rho(i, y) \quad \blacksquare
\end{aligned}
$$

Voters veto clockwise from top, vetoed alternative shown in red. Outcome is $x^\star = b$ and its social cost is compared to the optimum, d.

DISTORTION OF VOTING RULES

Rule	Metric distortion
k-approval ($k \geq 2$)	Unbounded
Plurality, Borda count	$\Theta(m)$
Best positional scoring rule	$\Omega\left(\sqrt{\log m}\right)$
IRV	$O(\log m), \Omega\left(\sqrt{\log m}\right)$
Llull	5
PluralityVeto	3

RANDOMIZED RULES

- Can randomized rules achieve better distortion?
- The utilitarian distortion of the best randomized rule is $\Theta(\sqrt{m})$
- The metric distortion of the best randomized rule is between 2.112 and 2.137

BIBLIOGRAPHY

A. D. Procaccia and J. S. Rosenschein. The Distortion of Cardinal Preferences in Voting. CIA 2006.

E. Anshelevitch, O. Bhardwaj, E. Elkind, J. Postl, and P. Skowron. Approximating Optimal Social Choice Under Metric Preferences. Artificial Intelligence, 2018.

F. E. Kizilkaya and D. Kempe. Plurality Veto: A Simple Voting Rule Achieving Optimal Metric Distortion. IJCAI 2022.