Optimized Democracy

Spring 2023 | Lecture 3

Strategic Manipulation

Ariel Procaccia | Harvard University

REMINDER: THE VOTING MODEL

- Set of voters $N = \{1, \dots, n\}$ (assume $n \geq 2$)
- Set of alternatives A; denote $|A| = m$
- Each voter has a ranking $\sigma_i \in \mathcal{L}$ over the alternatives; $x \succ_{\sigma_i} y$ means that voter i prefers x to y
- A preference profile $\boldsymbol{\sigma} \in \mathcal{L}^n$ is a collection of all voters' rankings
- A social choice function is a function $f: \mathcal{L}^n \to A$

MANIPULATION

So far the voters were honest!

MANIPULATION

- Using Borda count
- Top profile: *b* wins
- Bottom profile: *a* wins
- By changing their vote, voter 3 achieves a better outcome!

1	2	3
b	*b*	*a*
a	*a*	*b*
c	*c*	*c*
d	*d*	*d*

1	2	3
b	*b*	*a*
a	*a*	*c*
c	*c*	*d*
d	*d*	*b*

Jean-Charles de Borda

1733–1799

“My rule is intended for honest men!”

STRATEGYPROOFNESS

- Denote $\boldsymbol{\sigma}_{-i} = (\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_n)$
- A social choice function f is strategyproof (SP) if a voter can never benefit from lying about their preferences:

$$\forall \boldsymbol{\sigma} \in \mathcal{L}^n, \forall i \in N, \forall \sigma_i' \in \mathcal{L}, f(\boldsymbol{\sigma}) \succcurlyeq_{\sigma_i} f\left(\sigma_i', \boldsymbol{\sigma}_{-i}\right)$$

Poll 1

Max m for which plurality is SP?

- $m = 2$
- $m = 3$
- $m = 4$
- $m = \infty$

THE G-S THEOREM

- Theorem [Gibbard 1973, Satterthwaite 1975]: Let $m \geq 3$, then a social choice function f is SP and onto A (any alternative can win) if and only if f is dictatorial
- In other words, any voting rule that is onto and nondictatorial is manipulable

Question

For $m \geq 3$, all common rules are onto and nondictatorial. What about SP and nondictatorial?

PROOF SKETCH OF G-S

- Lemmas (prove in Pset 1):
  - Strong monotonicity: If f is SP function, $\boldsymbol{\sigma}$ profile, $f(\boldsymbol{\sigma}) = a$, then $f(\boldsymbol{\sigma}') = a$ for all profiles $\boldsymbol{\sigma}'$ s.t. $\forall x \in A, i \in N$: $\left[a \succ_{\sigma_i} x \Rightarrow a \succ_{\sigma_i'} x\right]$
  - Unanimity: If f is SP and onto function, $\boldsymbol{\sigma}$ profile, then $\left[\forall i \in N, a \succ_{\sigma_i} b\right] \Rightarrow f(\boldsymbol{\sigma}) \neq b$
- Let us assume that $m \geq n$, and neutrality:

$$f\big(\pi(\boldsymbol{\sigma})\big) = \pi\big(f(\boldsymbol{\sigma})\big) \text{ for all } \pi: A \to A$$

PROOF SKETCH OF G-S

- Say $n = 4$ and $A = \{a, b, c, d, e\}$
- Consider the following profile

$\boldsymbol{\sigma} =$

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

- Unanimity $\Rightarrow e$ is not the winner
- Suppose $f(\boldsymbol{\sigma}) = a$

PROOF SKETCH OF G-S

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

$\boldsymbol{\sigma}$

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

$\boldsymbol{\sigma}^1$

- Strong monotonicity $\Rightarrow f(\boldsymbol{\sigma}^1) = a$

PROOF SKETCH OF G-S

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

$\boldsymbol{\sigma}^1$

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

$\boldsymbol{\sigma}^2$

Poll 2

How many options are there for $f(\boldsymbol{\sigma}^2)$?

- 1 option
- 2 options
- 3 options
- 4 options

PROOF SKETCH OF G-S

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

$\boldsymbol{\sigma}^2$

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

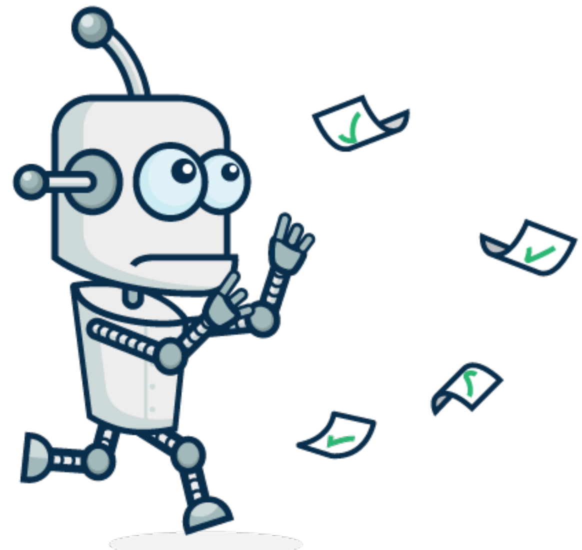
$\boldsymbol{\sigma}^3$

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

$\boldsymbol{\sigma}^4$

- Unanimity $\Rightarrow f(\boldsymbol{\sigma}^j) \notin \{b, c, e\}$
- $[\text{SP} \Rightarrow f(\boldsymbol{\sigma}^j) \neq d] \Rightarrow f(\boldsymbol{\sigma}^j) = a$
- Strong monotonicity $\Rightarrow f(\boldsymbol{\sigma}) = a$ for every $\boldsymbol{\sigma}$ where 1 ranks a first
- Neutrality $\Rightarrow$ 1 is a dictator $\blacksquare$

HARDNESS OF MANIPULATION



Manipulation may be unavoidable in theory, but we can we design “reasonable” voting rules where manipulation is computationally hard?

THE COMPUTATIONAL PROBLEM

- f-MANIPULATION problem:
  - Given votes of nonmanipulators and a preferred alternative p
  - Can manipulator cast vote that makes p uniquely win under f?
- Example: Borda, $p = a$

1	2	3
b	b	
a	a	
c	c	
d	d	

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

A GREEDY ALGORITHM

- Rank p in first place
- While there are unranked alternatives:
  - If there is an alternative that can be placed in next spot without preventing p from winning, place this alternative
  - Otherwise return false

EXAMPLE: BORDA

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

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

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

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

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

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

EXAMPLE: LLULL

1	2	3	4	5
a	*b*	*e*	*e*	*a*
b	*a*	*c*	*c*	
c	*d*	*b*	*b*	
d	*e*	*a*	*a*	
e	*c*	*d*	*d*	

Preference profile

	a	*b*	*c*	*d*	*e*
a	-	2	3	5	3
b	3	-	2	4	2
c	2	2	-	3	1
d	0	0	1	-	2
e	2	2	3	2	-

Pairwise comparisons

EXAMPLE: LLULL

1	2	3	4	5
a	*b*	*e*	*e*	*a*
b	*a*	*c*	*c*	*c*
c	*d*	*b*	*b*	
d	*e*	*a*	*a*	
e	*c*	*d*	*d*	

Preference profile

	a	*b*	*c*	*d*	*e*
a	-	2	3	5	3
b	3	-	2	4	2
c	2	3	-	4	2
d	0	0	1	-	2
e	2	2	3	2	-

Pairwise comparisons

EXAMPLE: LLULL

1	2	3	4	5
a	*b*	*e*	*e*	*a*
b	*a*	*c*	*c*	*c*
c	*d*	*b*	*b*	*d*
d	*e*	*a*	*a*	
e	*c*	*d*	*d*	

Preference profile

	a	*b*	*c*	*d*	*e*
a	-	2	3	5	3
b	3	-	2	4	2
c	2	3	-	4	2
d	0	1	1	-	3
e	2	2	3	2	-

Pairwise comparisons

EXAMPLE: LLULL

1	2	3	4	5
a	*b*	*e*	*e*	*a*
b	*a*	*c*	*c*	*c*
c	*d*	*b*	*b*	*d*
d	*e*	*a*	*a*	*e*
e	*c*	*d*	*d*	

Preference profile

	a	*b*	*c*	*d*	*e*
a	-	2	3	5	3
b	3	-	2	4	2
c	2	3	-	4	2
d	0	1	1	-	3
e	2	3	3	2	-

Pairwise comparisons

EXAMPLE: LLULL

1	2	3	4	5
a	b	e	e	a
b	a	c	c	c
c	d	b	b	d
d	e	a	a	e
e	c	d	d	b

Preference profile

	a	b	c	d	e
a	-	2	3	5	3
b	3	-	2	4	2
c	2	3	-	4	2
d	0	1	1	-	3
e	2	3	3	2	-

Pairwise comparisons

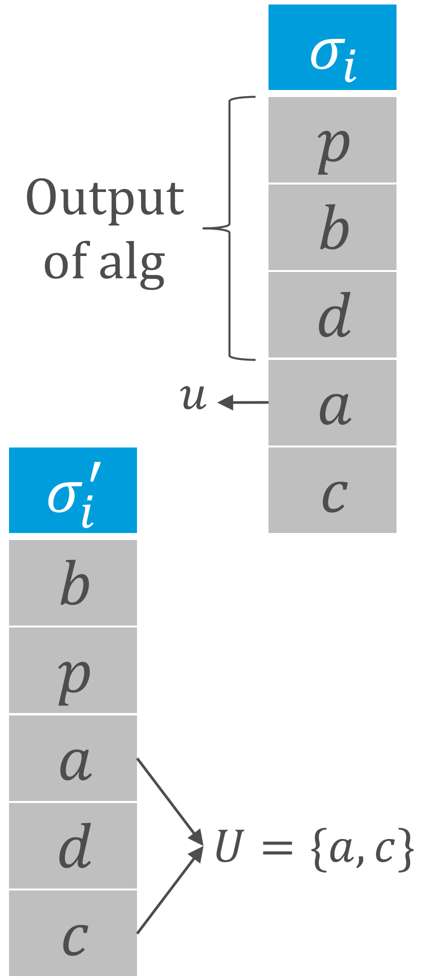
WHEN DOES THE ALG WORK?

- Theorem: Fix $i \in N$ and the votes of other voters. Let f be a rule s.t. $\exists$function $s(\sigma_i, x)$ such that:
  1. For every σ_i, f chooses an alternative that uniquely maximizes $s(\sigma_i, x)$
  2. If $\{y\colon\ y \prec_{\sigma_i} x\} \subseteq \{y\colon\ y \prec_{\sigma_i'} x\}$ then $s(\sigma_i, x) \le s(\sigma_i', x)$

  Then the greedy algorithm decides the f-MANIPULATION problem correctly

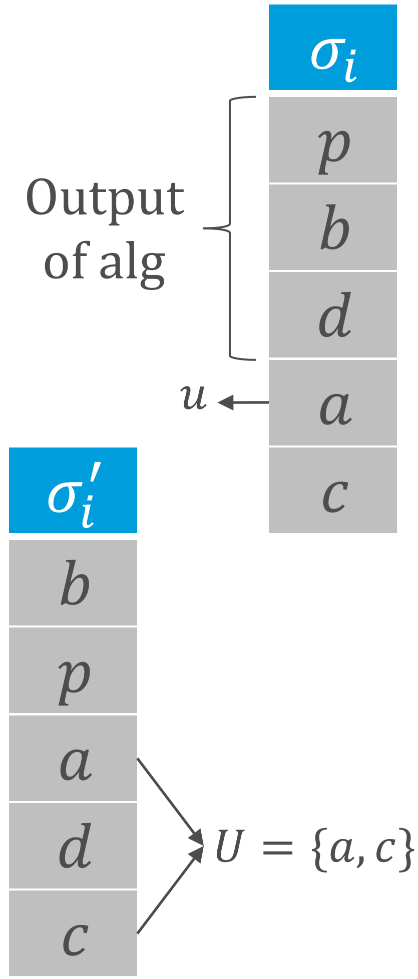
PROOF OF THEOREM

- Suppose the algorithm failed, producing a partial ranking σ_i
- Assume for contradiction σ_i' makes p win
- $U \leftarrow$ alternatives not ranked in σ_i
- $u \leftarrow$ highest ranked alternative in U according to σ_i'
- Complete σ_i by adding u first, then others arbitrarily

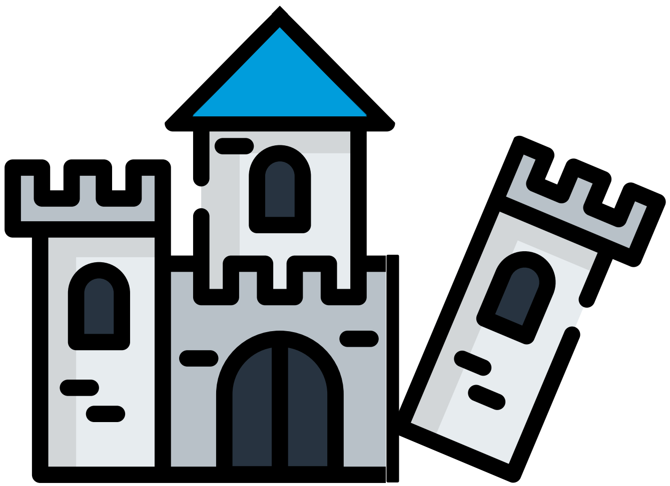


PROOF OF THEOREM

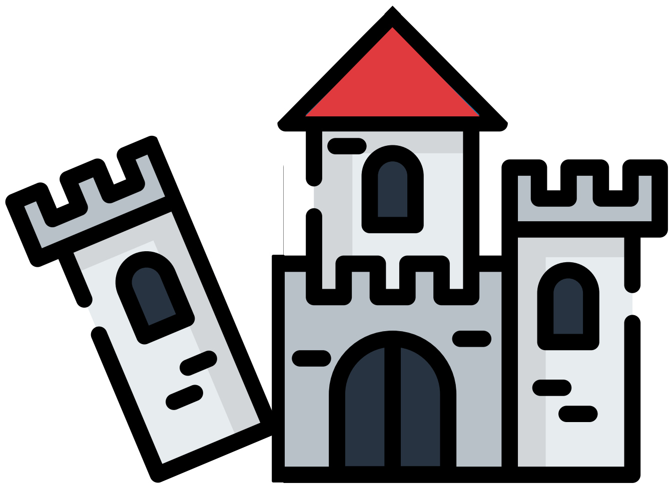
- Property 2 $\Rightarrow s(\sigma_i, p) \geq s(\sigma_i', p)$
- Property 1 and σ_i' makes p the winner $\Rightarrow s(\sigma_i', p) > s(\sigma_i', u)$
- Property 2 $\Rightarrow s\left(\sigma_i', u\right) \geq s(\sigma_i, u)$
- Conclusion: $s(\sigma_i, p) > s(\sigma_i, u)$, so the alg could have inserted u next ■



HARD-TO-MANIPULATE RULES



Single Transferable Vote



Llull (w. tie breaking)

But worst-case hardness isn't necessarily an obstacle to manipulation in the average case!

BIBLIOGRAPHY

A. Gibbard. Manipulation of Voting Schemes: A General Result. Econometrica, 1973.

M. A. Satterthwaite. Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions. Journal of Economic Theory, 1975.

J. Bartholdi, III, C. A. Tovey, and M. A. Trick. The Computational Difficulty of Manipulating an Election. Social Choice and Welfare, 1989.