TRUTH JUSTICE ALGOS

Social Choice III: Strategic Manipulation

Teachers: Ariel Procaccia (this time) and Alex Psomas

REMINDER: THE VOTING MODEL

- Set of voters $N = \{1, \dots, n\}$
- Set of alternatives A; denote $|A| = m$
- Each voter has a ranking $\sigma_i \in \mathcal{L}$ over the alternatives; $x \succ_i y$ 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 voting rule 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 his 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

STRATEGYPROOFNESS

- Denote $\boldsymbol{\sigma}_{-i} = (\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_n)$
- A voting rule is strategyproof (SP) if a voter can never benefit from lying about his preferences:
  $\forall \boldsymbol{\sigma} \in \mathcal{L}^n, \forall i \in N, \forall \sigma_i' \in \mathcal{L}, f(\boldsymbol{\sigma}) \succcurlyeq_i f(\sigma_i', \boldsymbol{\sigma}_{-i})$

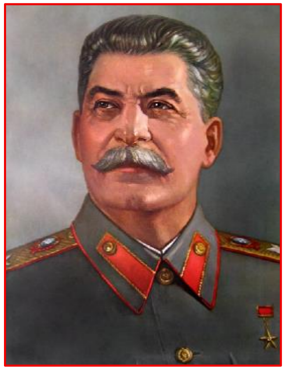
Question

Max m for which plurality is SP?

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

STRATEGYPROOFNESS

- A voting rule is **dictatorial** if there is a voter who always gets his most preferred alternative
- A voting rule is **constant** if the same alternative is always chosen
- Constant functions and dictatorships are SP



Dictatorship

Constant function

GIBBARD-SATTERTHWAITE

- A voting rule is onto if any alternative can win
- Theorem (Gibbard-Satterthwaite): If $m \geq 3$ then any voting rule that is SP and onto is dictatorial
- In other words, any voting rule that is onto and nondictatorial is manipulable

Gibbard

Satterthwaite

PROOF SKETCH OF G-S

- Lemmas (prove in HW1):
  - Strong monotonicity: f is SP rule, $\boldsymbol{\sigma}$ profile, $f(\boldsymbol{\sigma}) = a$. Then $f(\boldsymbol{\sigma}') = a$ for all profiles $\boldsymbol{\sigma}'$ s.t. $\forall x \in A, i \in N\colon\ [a \succ_i x \Rightarrow a \succ_i' x]$
  - Pareto optimality: f is SP+onto rule, $\boldsymbol{\sigma}$ profile. If $a \succ_i b$ for all $i \in N$ then $f(\boldsymbol{\sigma}) \neq b$
- Let us assume that $m \geq n$, and neutrality:

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

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

- Pareto optimality $\Rightarrow e$ 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 1

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

$\boldsymbol{\sigma}^3$

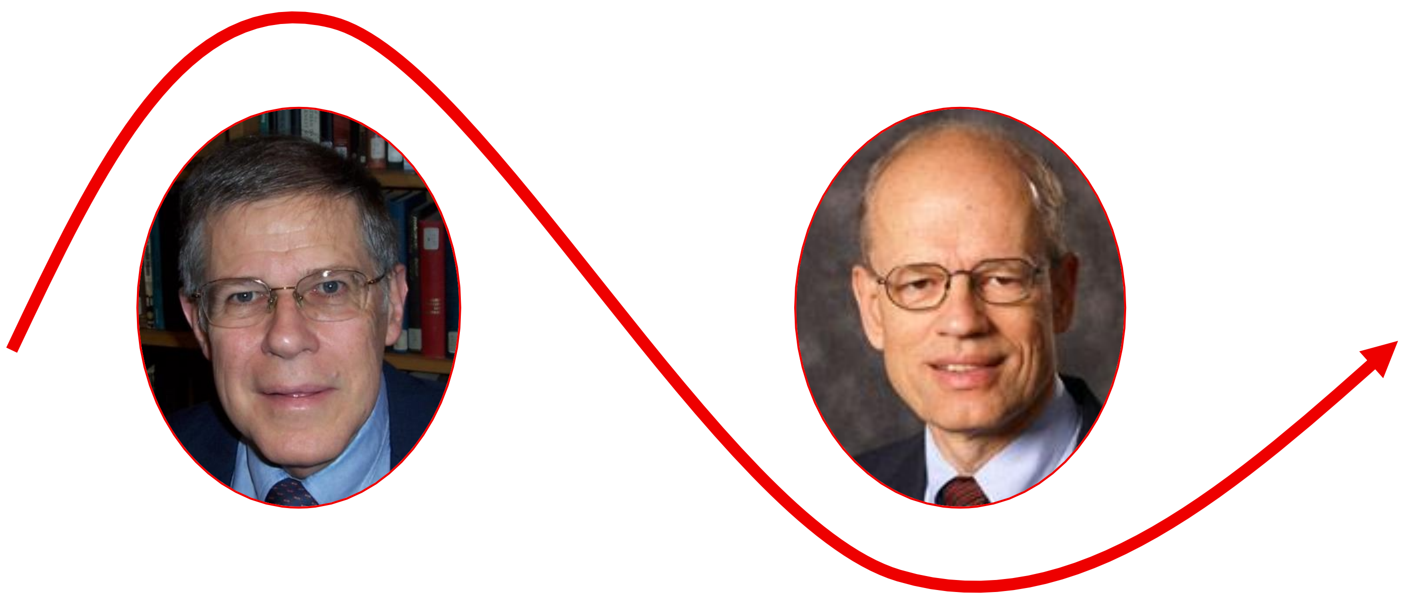
1	2	3	4
a	d	d	d
d	b	b	b
b	c	c	c
c	e	e	e
e	a	a	a

$\boldsymbol{\sigma}^4$

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

CIRCUMVENTING G-S

- Restricted preferences (next lecture)
- Money ⇒ mechanism design (done)
- Computational complexity (this lecture)



COMPLEXITY OF MANIPULATION

- Manipulation is always possible in theory
- But can we design voting rules where it is difficult in practice?
- Are there “reasonable” voting rules where manipulation is a hard computational problem? [Bartholdi et al. 1989]

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: COPELAND

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 elections

EXAMPLE: COPELAND

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 elections

EXAMPLE: COPELAND

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 elections

EXAMPLE: COPELAND

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 elections

EXAMPLE: COPELAND

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 elections

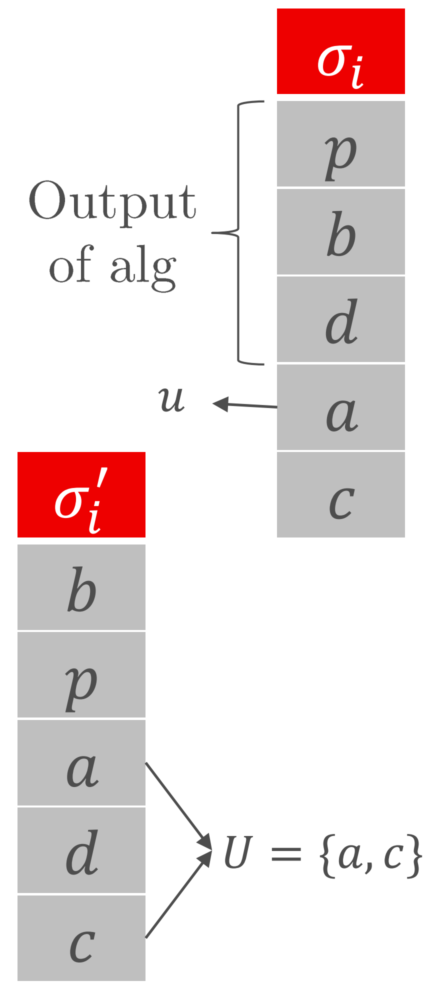
WHEN DOES THE ALG WORK?

- Theorem [Bartholdi et al., SCW 89]:
  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 a alternative that uniquely maximizes $s(\sigma_i, x)$
  2. $\{y\colon\ y \prec_i x\} \subseteq \{y\colon\ y \prec_i' x\} \Rightarrow s(\sigma_i, x) \leq s(\sigma_i', x)$

  Then the algorithm always decides f-MANIPULATION correctly

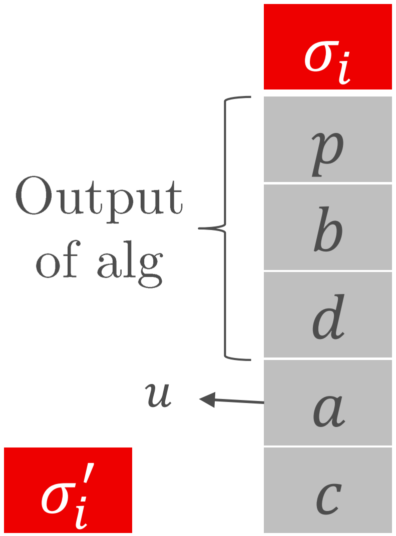
PROOF OF THEOREM

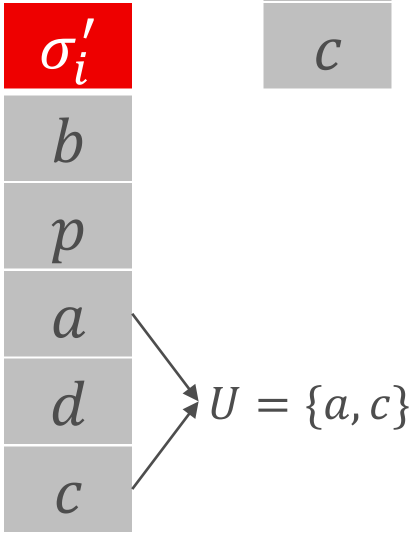
- Suppose the algorithm failed, producing a partial ranking σ_i
- Assume for contradiction σ_i' makes p win
- $U \leftarrow$ alternatives not ranked in σ_i
- $u \leftarrow$ highest ranked alternative in U according to σ_i'
- Complete σ_{i} by adding u first, then others arbitrarily



PROOF OF THEOREM

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

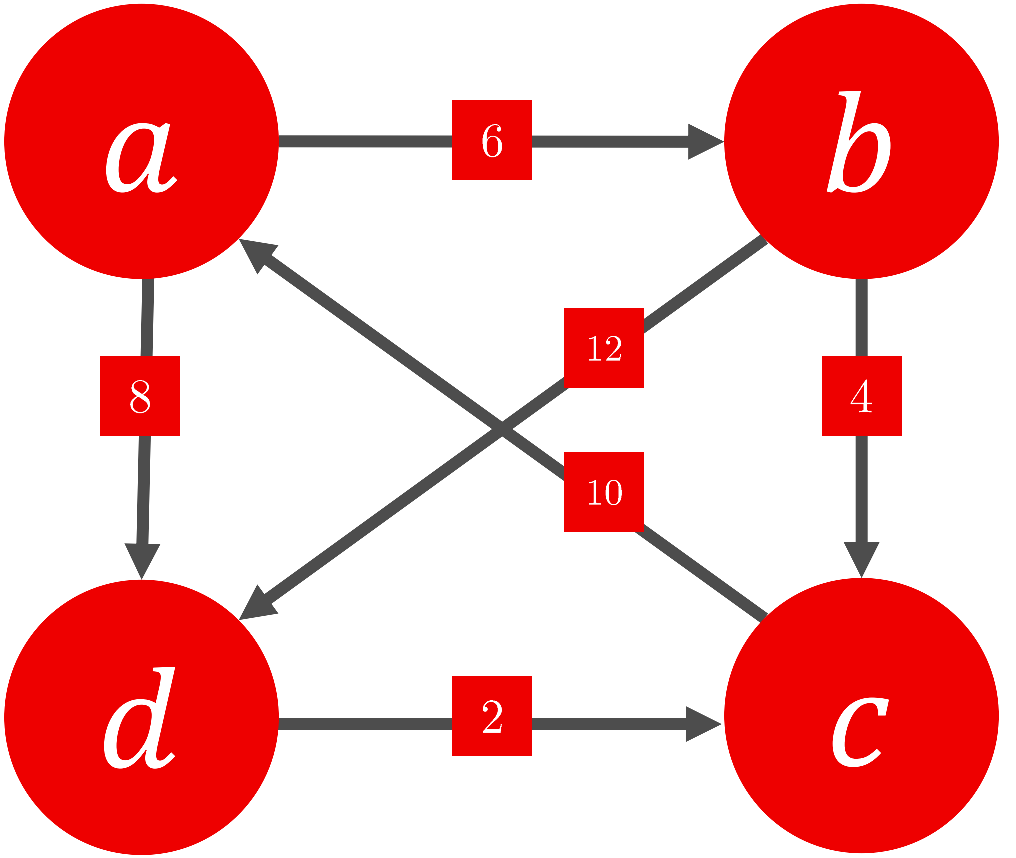




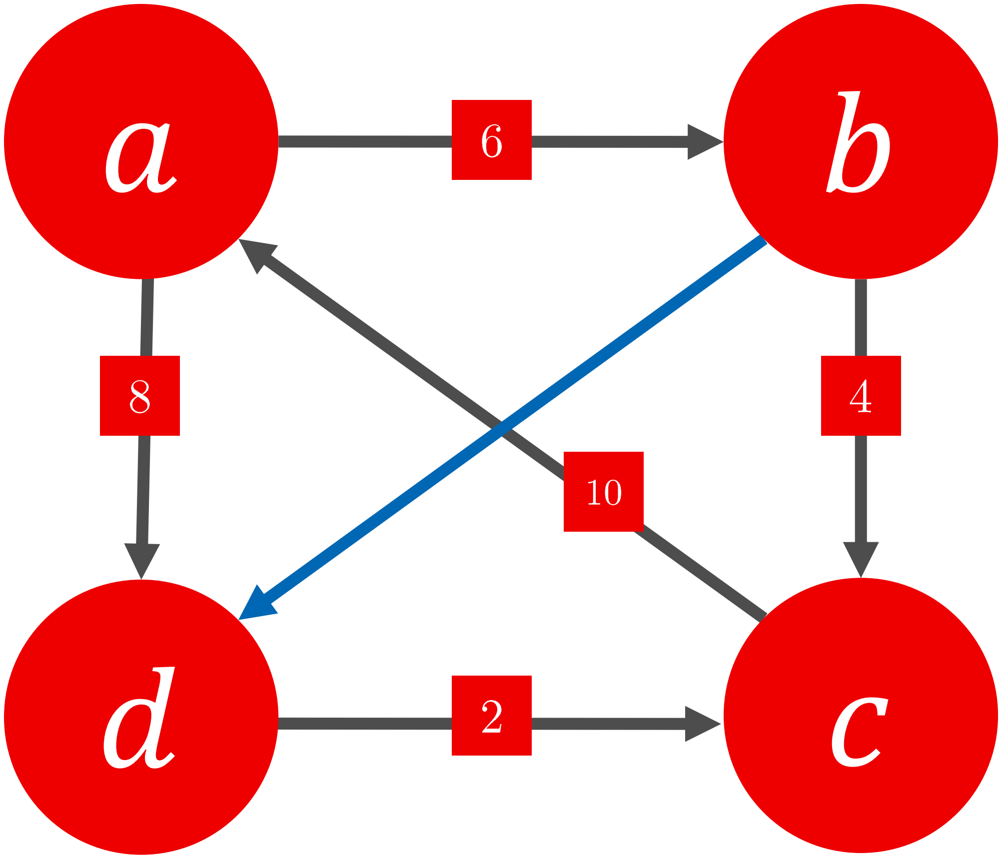
HARD-TO-MANIPULATE RULES

- Copeland with second order tie breaking [Bartholdi et al. 1989]
- STV [Bartholdi and Orlin 1991]
- Ranked Pairs [Xia et al. 2009]
  - Sort pairwise comparisons by strength
  - Lock in pairwise comparisons in that order, unless a cycle is created, in which case the opposite edge is locked in
  - Return the alternative at the top of the induced order

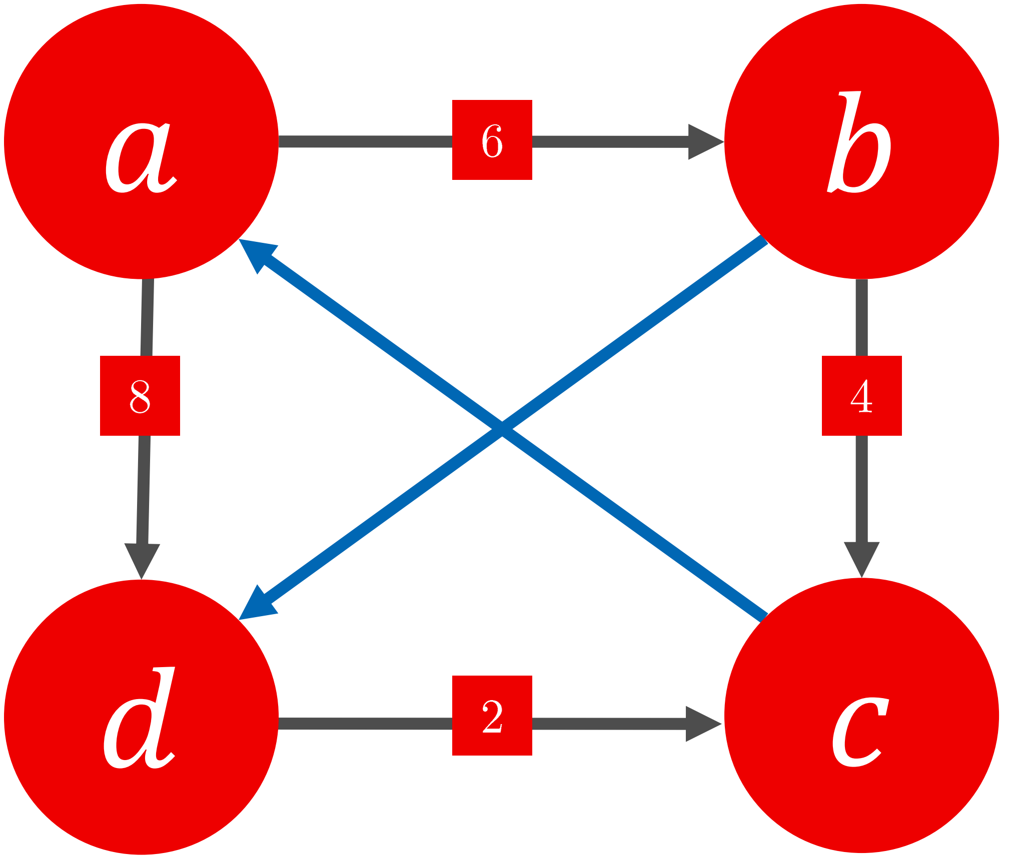
EXAMPLE: RANKED PAIRS



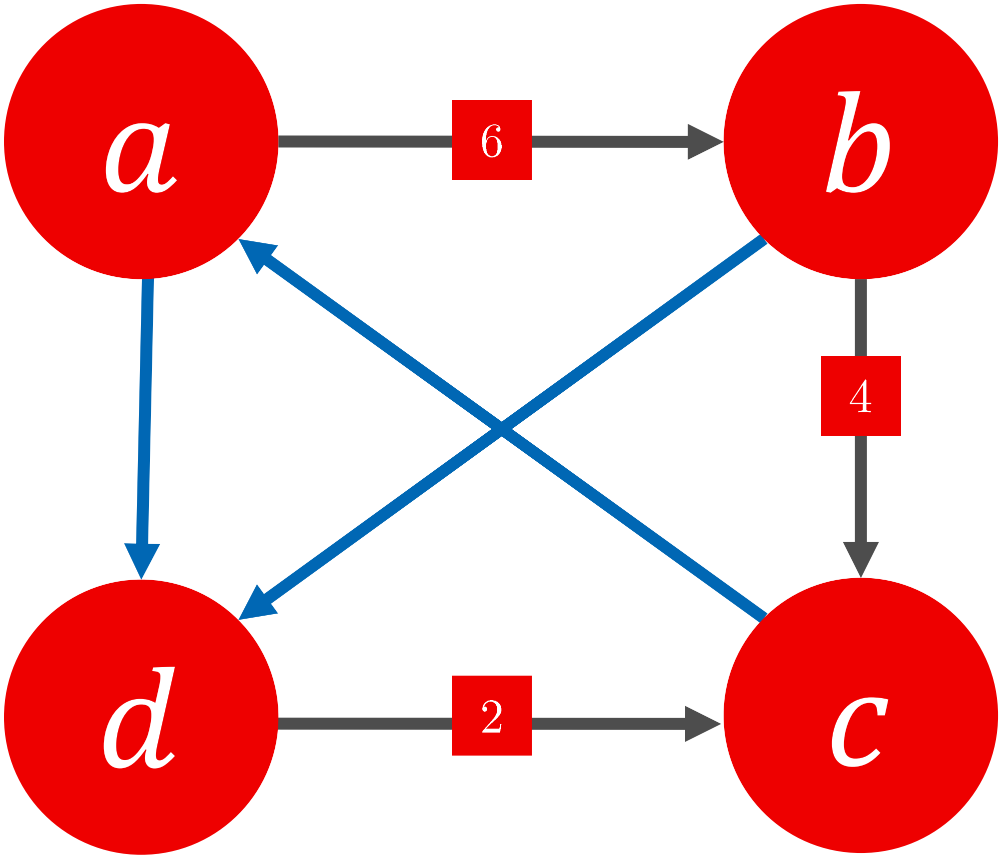
EXAMPLE: RANKED PAIRS



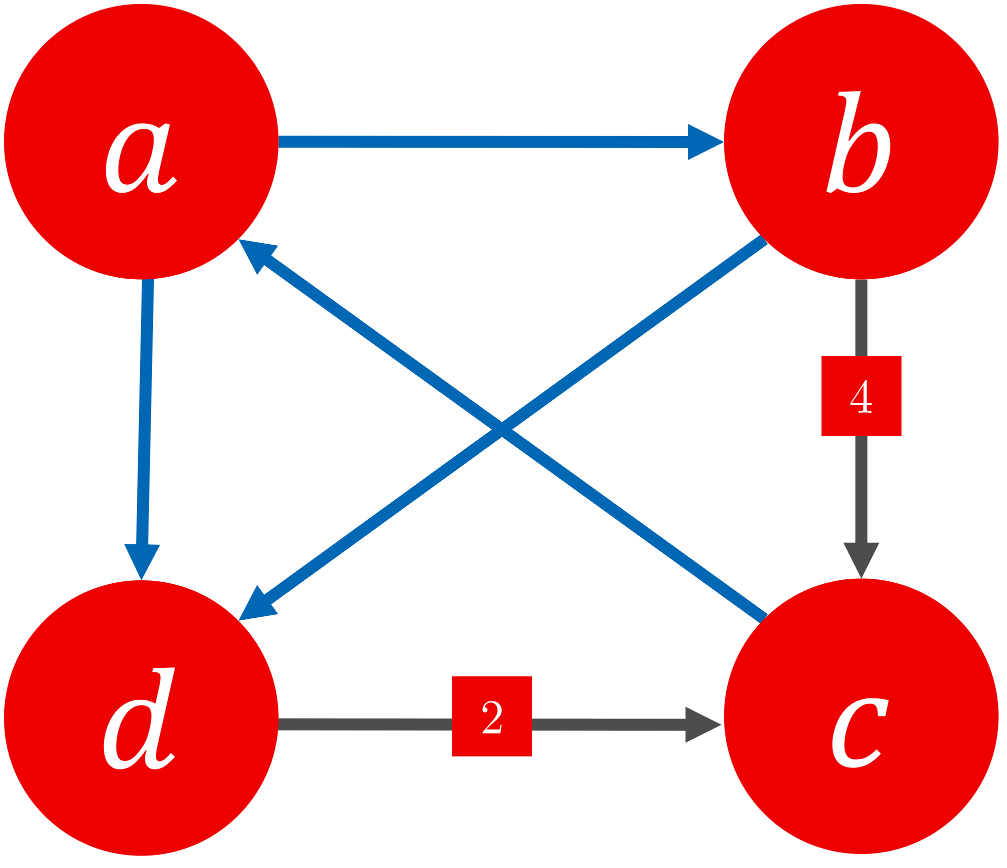
EXAMPLE: RANKED PAIRS



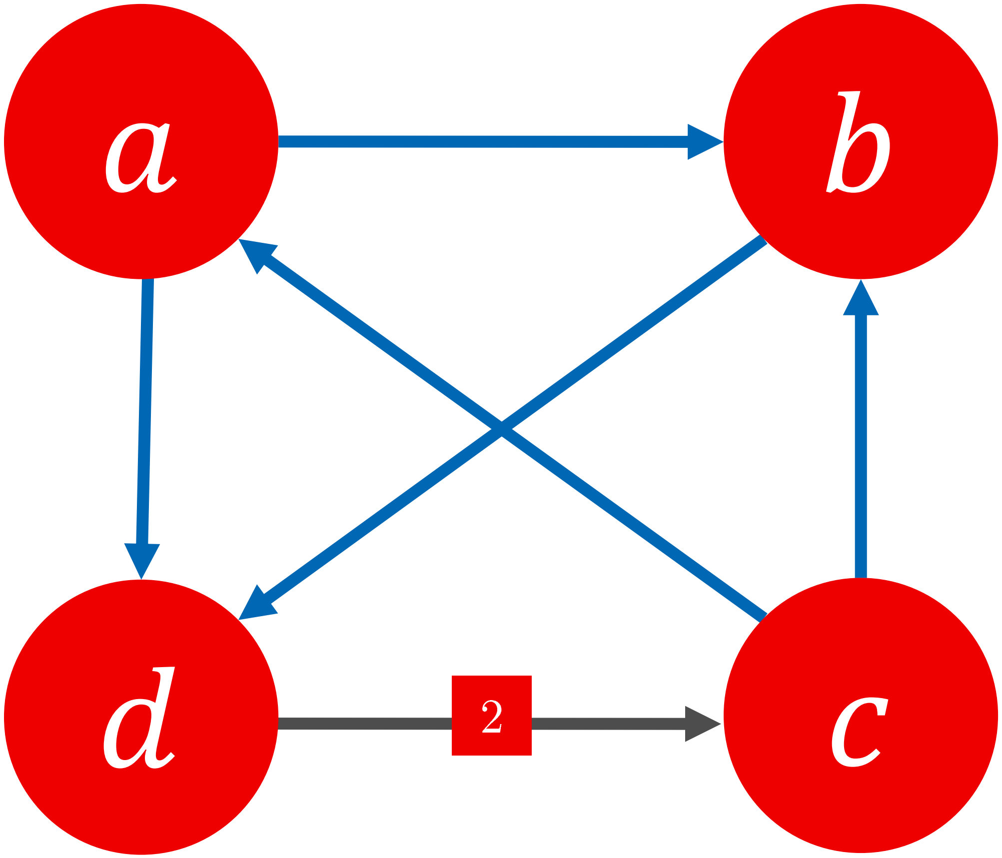
EXAMPLE: RANKED PAIRS



EXAMPLE: RANKED PAIRS



EXAMPLE: RANKED PAIRS



EXAMPLE: RANKED PAIRS