Algorithmic Game Theory

Lecture 4

Jiaxin Song

Recall: Strategy-Proofness

- A social choice function f can be strategically manipulated by voter i if, for some $<_1, \ldots, <_n \in L$ and some $<_i' \in L$, we have $a <_i a'$ where $\alpha = f(<_1, \ldots, <_i, \ldots, <_n)$ and $\alpha' = f(<_1, \ldots, <_i', \ldots, <_n)$.
- f is strategy-proof if it is not strategically manipulatable by any agent.
- Strategy-Proof/Truthful/Incentive-Compatible:
    - "truth-telling is a dominant strategy"

Preliminaries

- An allocation rule is a function $f : \times \mathscr{R}_i \to A$, mapping preferences of the agents into alternatives.
- The *allocation space* is the unit interval $A = [0,1]$.
- An *outcome* in this model is a single point $x \in A$.
- Each agent $i \in N$ has a preference ordering $\succeq_i$ over the outcomes in $[0,1]$.
- The preference relation $\succeq_i$ is *single-peaked* if there exists a point $p_i \in A$ (the *peak* of $\succeq_i$) such that for all $x \in A \backslash \{p_i\}$ and all $\lambda \in [0,1)$, $\lambda x + (1 - \lambda)p_i \succ_i x$
- Let $\mathscr{R}$ denote the class of single-peaked preferences.
- We denote the peaks of preference relations $\succeq_i$, $\succeq_i'$ by p_i, p_i'.

Single-Peaked Preferences

- ***Theorem.*** *Suppose f is **strategy-poof**. Then f is **onto** $\Leftrightarrow$ it is **unanimous** $\Leftrightarrow$ if it is **Pareto-optimal**.*
  - ***Proof.***

1) Strategy-proof + onto $\Rightarrow$ unanimous

Fix $x \in [0,1]$, and consider any unanimous profile $\succeq \in \mathscr{R}^n$ such that $p_i = x, \forall i \in N$.

Let $\succeq' \in \mathscr{R}^n$ be such that $f(\succeq') = x$.

By strategy-proofness, $f(\succeq_1, \succeq_2', \ldots, \succeq_n') = x$,

Otherwise, if agent 1's true preference is $\succeq_1$, (s)he could misreports his(her) valuation as $\succeq_1'$ and manipulate the outcome

Repeating this argument, $f(\succeq_1, \succeq_2, \ldots, \succeq_n) = x$

Single-Peaked Preferences

- ***Theorem.*** *Suppose f is **strategy-poof**. Then f is **onto** $\Leftrightarrow$ it is **unanimous** $\Leftrightarrow$ if it is **Pareto-optimal**.*
  - ***Proof.***

1) **Strategy-proof + unanimous $\Rightarrow$ Pareto-optimal**

Suppose f is not Pareto-optimal at some profile $\succeq \in \mathscr{R}^n$.

There must exist $x \in [0,1]$ such that $x \succeq_j f(\succeq), \forall j \in N$ and there exists $i \in N$ such that $x \succ_i f(\succeq)$.

Assume $p_1 \leq p_2 \leq \ldots \leq p_n$

The above fact implies $f(\succeq) < p_j, \forall j \in N$.

a) If $p_1 = p_n$, it violates the unanimity since all peaks are equal but $p_1 > f(\succeq)$

Single-Peaked Preferences

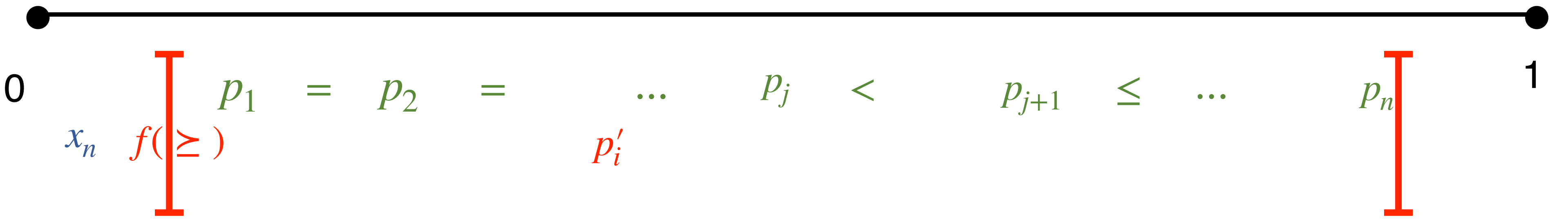
- ***Theorem.*** *Suppose f is* ***strategy-poof****. Then f is* ***onto*** *⇔ it is* ***unanimous*** *⇔ if it is* ***Pareto-optimal****.*

b) If $p_1 = \ldots = p_j < p_{j+1} \leq \ldots \leq p_n$,

For all $i > j$, let $\succeq_i'$ be a preference relation such that both $p_i' = p_1$ and $f(\succeq) \succeq_i' p_i$.

Let $x_n = f(\succeq_1, \ldots, \succeq_{n-1}, \succeq_n')$.

By *strategy-proofness*, $x_n \in [f(\succeq), p_n]$, otherwise agent n with preference $\succeq_n'$ could manipulate f by reporting preference $\succeq_n$

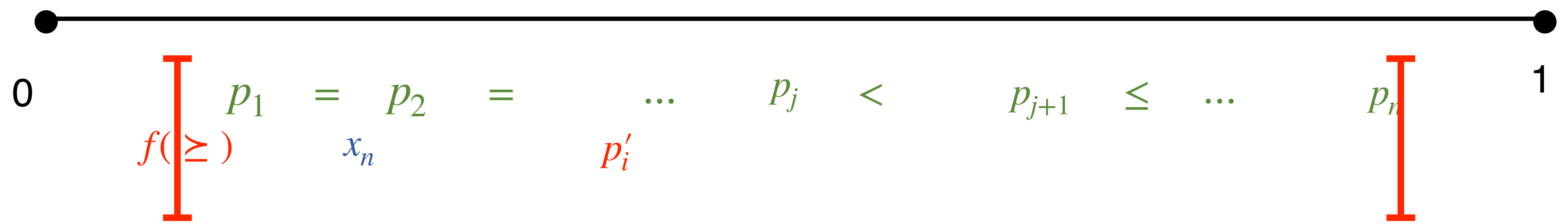


Single-Peaked Preferences

- ***Theorem.*** *Suppose f is **strategy-poof**. Then f is **onto** ⇔ it is **unanimous** ⇔ if it is **Pareto-optimal**.*

b) If $p_1 = \ldots = p_j < p_{j+1} \leq \ldots \leq p_n$,

Similarly, $x_n \notin (f(\succeq), p_n]$, otherwise agent n with preference $\succeq_n$ could manipulate f by reporting preference $\succeq_n'$



Rules

- Could you find a *strategy-proof* strategy on this domain?
  - Choose agent 1's peak,
  - Just choose a constant value belonging to [0,1].
- (*Median-voter Rule*) Suppose that the number of agents n is odd, then picks the median of the agents' peaks (p_i's)
- The above rule is *strategy-proof*:
  - If an agent's peak p_i lies *below* the median peak, then he can only change the output by reporting a preference relation whose peak lies *above* the true median.
  - This effect of this misreport is for the rule to choose a point even further away from p_i, making the agent worse off.

Rules

- More generally, for any number of agents n and any positive integer $k \leq n$, the rule that picks the k-th highest peak is *strategy-proof.*
- If we choose *average* rather than *medium*?
- (❓ Question 1) Is the rule of choosing the **average** of the n agents's peaks *strategy-proof* ?
- (❓ Question 2) Is the rule of choosing the **weighted average** of the n agents's peaks *strategy-proof* ?

Rules

Theorem. *A rule f is **strategy-poof, onto** and **anonymous** ⇔ if there exist* $y_1, y_2, \ldots, y_{n-1} \in [0,1]$ *such that for all* $\succeq \in \mathcal{R}^n$,

$$f(\succeq) = med\{p_1, p_2, \ldots, p_n, y_1, y_2, \ldots, y_{n-1}\}$$

Sketch of Proof.

1) ***Sufficiency.*** Easy to verify!

2) ***Necessity.***

Since preferences relations are ordinal, there is only one preference relation with a peak at 0 and only with a peak at 1. Denote them as $\succeq_i^0$ and $\succeq_i^1$ respectively.

For any $1 \leq m \leq n-1$, let y_m be the following outcome:

$$y_m = f(\succeq_1^0, \ldots, \succeq_{n-m}^0, \succeq_{n-m+1}^1, \ldots, \succeq_n^1)$$

Rules

For any $1 \leq m \leq n-1$, let y_m be the following outcome

$$y_m = f(\succeq_1^0, \ldots, \succeq_{n-m}^0, \succeq_{n-m+1}^1, \ldots, \succeq_n^1)$$

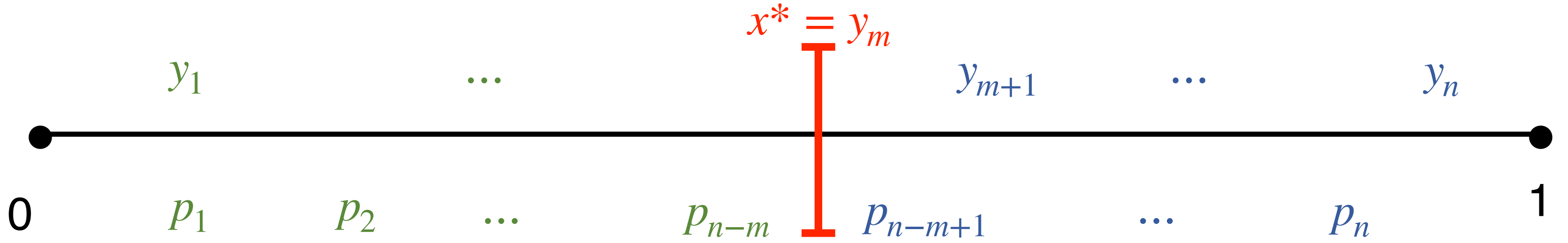
According to the *strategy-proofness*,

$$y_m \leq y_{m+1}, \forall 1 \leq m \leq n-2.$$

Consider a profile of preference $\succeq \in \mathcal{R}^n$ with peaks $p_1, \ldots, p_n$. Assume $p_i \leq p_{i+1}, \forall i \leq n-1$, we wish to show that:

$$f(\succeq) = x^* \equiv med\{p_1, \ldots, p_n, y_1, \ldots, y_{n-1}\}$$

Rules



(Case 1: the median is some y_m.)

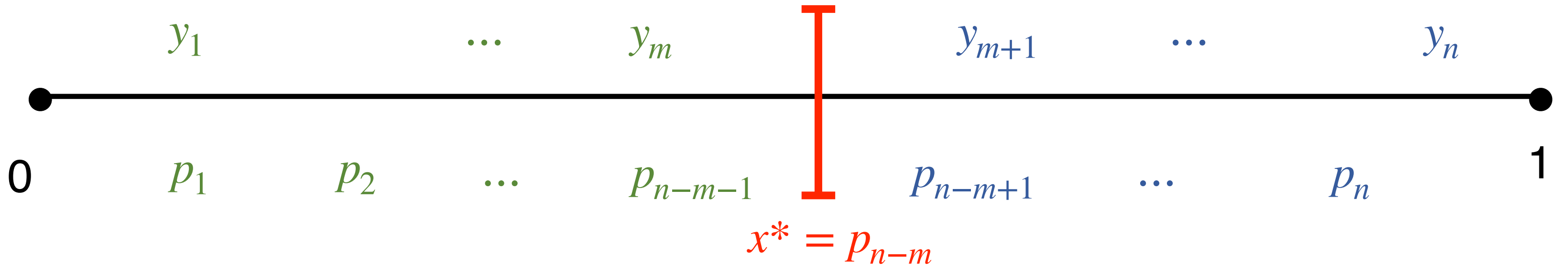
Suppose $x^* = y_m$ for some m, since x^* is the median of all the $2n-1$ points, this implies $p_{n-m} \leq x^* = y_m \leq p_{n-m+1}$. By assumption,

$$x^* = y_m = f(\succeq_1^0, \ldots, \succ_{n-m}^0, \succeq_{n-m+1}^1, \ldots, \succeq_n^1)$$

Let $x_1 = f(\succeq_1, \succeq_2^0 \ldots, \succ_{n-m}^0, \succeq_{n-m+1}^1, \ldots, \succeq_n^1)$, then we argue that $x^* = x_1$:

- If $x^* < x_1$, then agent 1 with preference $\succeq_1$ could manipulate f.
- If $x^* > x_1$, then agent 1 with preference $\succeq_1^0$ could manipulate f.

Rules



(Case 2: the median is an agent's peak .)

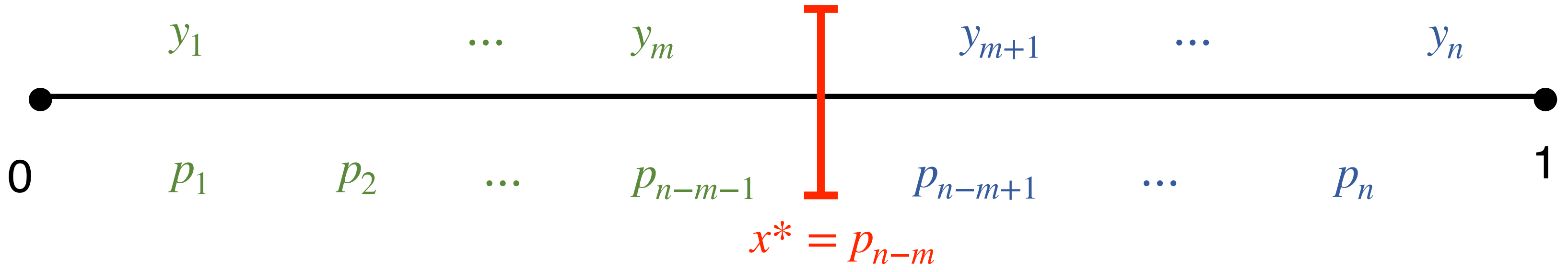
If $y_m < x^* < y_{m+1}$ for some m, we have $x^* = p_{n-m}$.

We can discuss the comparison between

$f(\succeq_1^0, \ldots, \succeq_{n-m-1}^0, \succeq_{n-m}, \succeq_{n-m+1}^1, \ldots, \succeq_n^1)$ and x^*.

The full proof is omitted here!

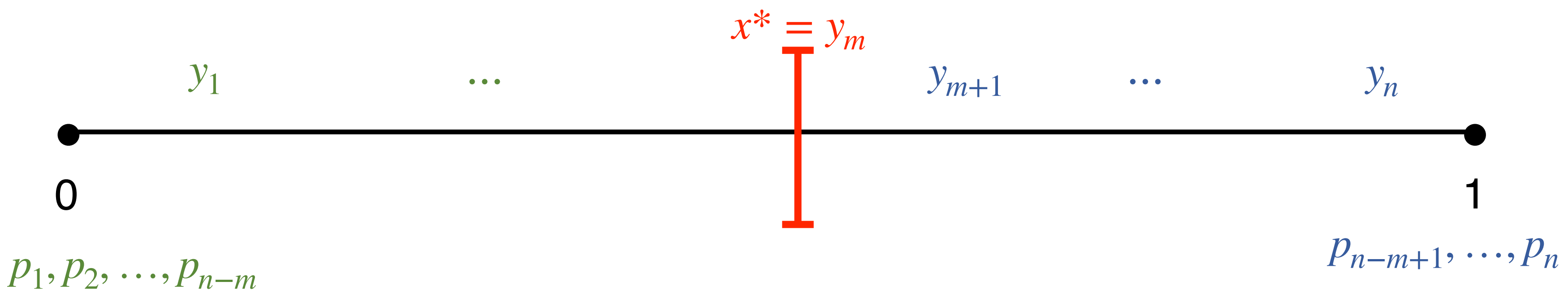
Rules



- If $x^* > x' = f(\succeq_1^0, \ldots, \succeq_{n-m-1}^0, \succeq_{n-m}, \succeq_{n-m+1}^1, \ldots, \succeq_n^1),$

Rules

- The parameters y_m can be thought of as the rule's degree of compromise when agents have extremist preferences.



- The above rules consider *anonymity*. But some some *strategy-proof* rules we have mentioned before are *dictatorial*. How to determine a more generalized *strategy-proof, onto* rule?

Rules

Definition. *A rule f is a generalized median voter scheme if there exist 2^n points in* $[0,1]$ *,* $\{\alpha_S\}_{S \subseteq N}$*, such that*

- $S \subseteq T \subseteq N$ implies $\alpha_S \leq \alpha_T$,
- $\alpha_\varnothing = 0, \alpha_N = 1$, and
- For all $\succeq \in \mathscr{R}^n, f(\succeq) = \max_{s \subset N} \min\{\alpha_S, p_i : i \in S\}$

- $S = \{\varnothing\}, \min\{\alpha_S, p_i : i \in S\} = \alpha_\varnothing,$
- $S = \{1\}, \min\{\alpha_S, p_i : i \in S\} = \alpha_1,$
- $S = \{2\}, \min\{\alpha_S, p_i : i \in S\} = \alpha_2,$
- $S = \{3\}, \min\{\alpha_S, p_i : i \in S\} = \alpha_3,$
- $S = \{1,2\}, \min\{\alpha_S, p_i : i \in S\} = \alpha_1,$
- $S = \{2,3\}, \min\{\alpha_S, p_i : i \in S\} = \alpha_2,$
- $S = \{1,3\}, \min\{\alpha_S, p_i : i \in S\} = \alpha_1,$
- $S = \{1,2,3\}, \min\{\alpha_S, p_i : i \in S\} = \alpha_1$

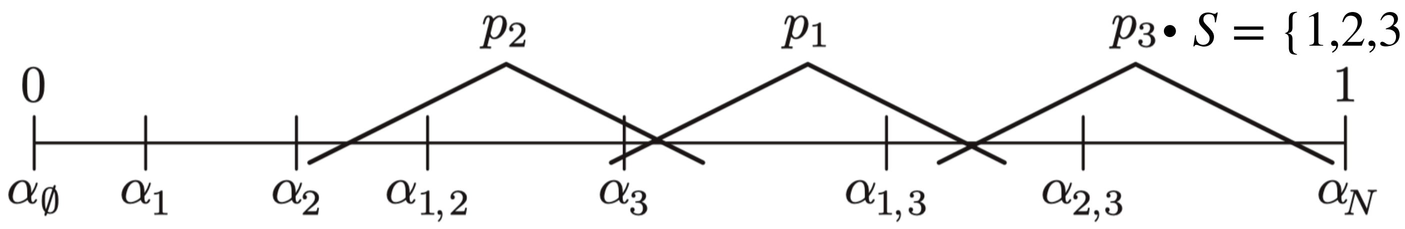


Figure 10.2. An example of a generalized median voter scheme for $n = 3$.

Rules

Definition. *A rule f is a generalized median voter scheme if there exist* 2^n *points in* $[0,1]$, $\{\alpha_S\}_{S\subseteq N}$, *such that*

- $S \subseteq T \subseteq N$ implies $\alpha_S \leq \alpha_T$,
- $\alpha_\varnothing = 0, \alpha_N = 1$, and
- For all $\succeq \in \mathscr{R}^n, f(\succeq) = \max_{s \subset N} \min\{\alpha_S, p_i : i \in S\}$

Theorem. *A rule f is* ***strategy-poof, onto*** $\Leftrightarrow$ *it is a generalized median voter scheme .*

Application 1: Truthful Cake Cutting

- Consider allocating a piece of cake to n agents, the cake is modeled as the interval [0,1].
- Each agent i has a *value density function* $f_i : [0,1] \to \mathbb{R}_{\geq 0}$ that describes his/her preference on the cake.
- A value density function $f_i : [0,1] \to \mathbb{R}_{\geq 0}$ is *piecewise-constant* if [0,1] can be partitioned into finitely many intervals, and f_i is constant on each of these interval.
- Given a subset $X \subseteq [0,1]$, agent i's *utility* on X, denoted by $v_i(X)$ is given by:

$$v_i(X) = \int_X f_i(x)dx$$

Application 1: Truthful Cake Cutting

- An allocation is *proportional* if each agent receives his/her average share of the entire cake:

$$\forall i : v_i(A_i) = \frac{1}{n} v_i([0,1])$$

- An allocation is *envy-free* if each agent receive a portion that has weakly higher value than any portion received by any other agent, based on his/her own valuation:

$$\forall i, j : v_i(A_i) \geq v_i(A_j)$$

- A mechanism M is *truthful* if each agent's dominant strategy is to report his/her true value density function. That is, for each $i \in [n]$, any $(f_1, \ldots, f_n)$ and any f_i',

$$v_i(\mathcal{M}_i(f_1, \ldots, f_n)) \geq v_i(\mathcal{M}_i(f_1, \ldots, f_{i-1}, f_i', f_{i+1}, \ldots, f_n))$$

Application 1: Truthful Cake Cutting

- Prof. Tao has given the following negative result in EC'22.

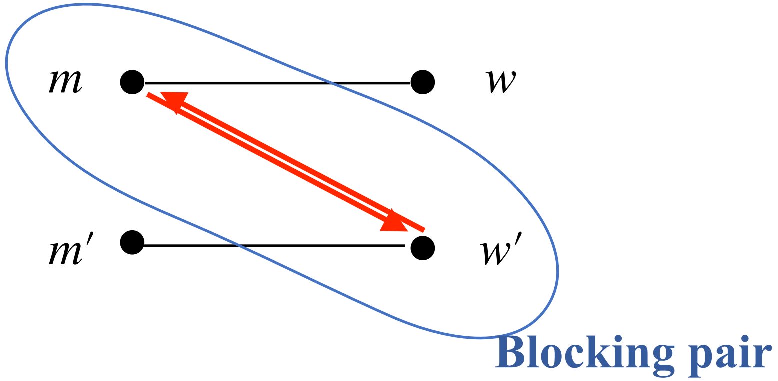
Theorem 3.1. *There does not exist a truthful proportional mechanism, even when all of the followings hold:*

- *there are two agents;*
- *each agent's value density function is piecewise-constant;*
- *each agent is hungry: each f_i satisfies $f_i(x) > 0$ for any $x \in [0,1]$;*
- *the mechanism needs not to be entire.*

- There are still three open problems:
  - Does there exist a positive integer $n \geq 3$ such that there exists a truthful proportional mechanism with n agents?
  - Does there exist an $\alpha > 0$ such that there exists a truthful, α-approximately proportional mechanism?
  - Does there exist a truthful mechanism that always allocates each agent a subset on which the agent has a positive value?

Application 3: Stable Matchings

- There are a set M of men and a set W of women.
- Each $m \in M$ has a strict preference ordering over the elements of W and each $w \in W$ has a strict preference ordering over the men.
- $x \succ_i y$ will mean that agent i ranks x above y
- We assume that $|M| = |W|$, and want to find a matching between the two sets.
- A matching is called **unstable** if there are two men m, m' and two women w, w' such that:
  - m is mathed to w
  - m' is matched to w', and
  - $w' \succ_m w$ and $m \succ_{w'} m'$



Male-Proposals

- First, each male proposes his top-ranked choice.
- Next, each woman who has received at least two proposals keeps her top-ranked proposal and rejects the rest.
- Then, each man who has been rejected proposes to his top-ranked choice among the women who have not rejected him
- ……
- **Theorem** *The male propose algorithm terminates in a stable matching.*

Male/Female-Optimal

- A matching μ is **male-optimal** if there is no stable matching v such that $v(m) \succ_m \mu(m)$ or $v(m) = \mu(m)$ for all m with $v(j) \succ_j \mu(j)$ for at least one $j \in M$
- **Theorem** *The male propose algorithm terminates in a stable matching.*
- *Proof.* Consider the first iteration such that $v(j) \succ_j \mu(j)$ first occurs.

The LP Formulation

- For each man m and woman w, let $x_{mw} = 1$ if man m is matched with woman w and zero otherwise.
- Then every stable matching must satisfy the following:

$$\sum_{w \in W} x_{mw} = 1, \forall m \in M$$

$$\sum_{m \in M} x_{mw} = 1, \forall w \in W$$

$$\sum_{j \prec_m w} x_{mj} + \sum_{i \prec_w m} x_{iw} + x_{mw} \leq 1, \forall m \in M, w \in W$$

$$x_{mw} \geq 0, \forall m \in M, w \in W$$

- *Proof.* Consider the first iteration such that $v(j) \succ_j \mu(j)$ first occurs.

This Lecture

- Social choice
  - Arrow's impossibility theorem
  - Gibbard-Satterthwaite Theorem
    - Strategy-proof is impossible
- Social choice with money
  - VCG mechanism (with Clarke Pivot Rule)
    - Maximizes social welfare
    - Enable strategy-proofness