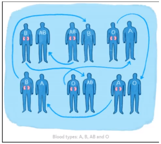
Two-sided matching theory

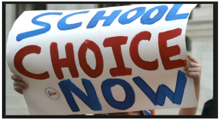
Somouaoga BONKOUNGOU

National Research University Higher School of Economics

March 4, 2019

- Medical match
- Allocation of dorms
- Assignment of cadets to branches
- Kidney exchange
- Yandex taxi?
- School choice
- Refugee resettlement
- Student exchange programs
- etc.





Characteristics

- Indivisibilities,
- two-sided (mostly),
- markets without prices,
- matching under preferences (our interest).

Outline

- Model
  - Formal model
  - Assumptions
- Stability
  - existence,
  - opposition of interest
  - lattice structure
- Incentive
  - impossibility
  - partial possibility

Benchmark model: marriage market

There are two finite and disjoint sets M and W:

- $M = \{m_1, m_2, \ldots, m_n\}$ is the set of men, and
- $W = \{w_1, w_2, \ldots, w_p\}$ is the set of women.

- There is a profile $P = (P_{m_1}, \ldots, P_{m_n}, P_{w_1}, \ldots, P_{w_p})$ of preferences where:
- for each man m, P_m is m's preference ordering over $W \cup \{m\}$
- for each woman w, P_w is w's preference ordering over $M \cup \{w\}$,
- a profile P is called a market.

Marriage market

Assumption

- *We assume that each man (woman) has strict preferences,*
- *agents have access to each other preferences.*

Matching

A matching $\mu : M \cup W \to M \cup W$ is a one-to-one function such that

- each man m is either matched to one woman or remains single: $\mu(m) = w$ or $\mu(m) = m$,
- each woman w is either matched to one man or remains single: $\mu(w) = m$ or $\mu(w) = w$,
- A man m is matched to women w if, and only if, woman w is matched to man m: $\mu(m) = w \Leftrightarrow \mu(w) = m$.

Which matchings are likely to occurs?

Rule

No man or woman is compelled to marry (consenting marriage).

Which matchings are likely to occur?

Observation

We will not observe any matchings which could only result from compulsion of one of the agents.

Prediction

Given that agents are rational,

- *any matching μ in which m and w are matched to each other and m is not acceptable to w will not occur.*
- *any matching μ such that there exists m and w who prefer each other to their mates at μ, will not occur:*

  m and w have a reason to disrupt μ and marry each other.

Which matchings are likely to occur?

Two criteria for excluding potential matchings:

Definition

- An individual blocks a matching if he prefers the option of being single to his mate,
- a pair (m, w) blocks a matching if they prefer each other to their mates,
- a matching is stable if it is not blocked by any individual or any pair of agents.

Stability: existence

Question: do stable matchings exist for each market?

One day, a letter arrived from Gale framing a problem of choosing roommates. If you have two groups, with each individual having different preferences, is there a way to come up with a set of stable pairings of one from each group? Gale suspected there wasn't a way to make a stable solution, but couldn't prove it.

Stability: existence

Question: do stable matchings exist for each market?

The way my father describes it, he received the letter around noon, spent some time thinking, wrote up the solution (There *was* a stable solution, and this is how you come up with it), and mailed it back to Gale later that afternoon.

Deferred acceptance

Deferred Acceptance algorithm

- Men propose women in the same order as in their preferences:
  - Each man will start proposing his most preferred woman
  - If rejected, a man will propose his second most preferred woman
  - If rejected, a man will propose his third most preferred woman
  - etc.
- Each woman always keeps the best man (according to her preferences) among the man proposing her (if any), and rejects the others.
- The algorithm stops when there's no more rejection.

Deferred Acceptance

Step 1

Each man proposes to his most preferred, acceptable woman

(if a man finds all women unacceptable he remains single).

Each woman who received at least one offer

- temporarily holds the offer from the most preferred man among those who made an offer to her and are acceptable.
- rejects the other offer(s).

Step $\mathbf{k}, \mathbf{k} \geq 2$

Each man whose offer has been rejected in the previous step proposes to his most preferred woman among the acceptable women he has not yet proposed.

(if there is no such woman he remains single).

Each woman who received at least one offer in this step

- temporarily holds the offer from the most preferred man among
  - those who made an offer to her in this step and are acceptable.
  - the man she held from the previous step (if any).

- rejects the other offer(s).

End: The algorithm stops when no man has an offer that is rejected.

Final matching:

- Each woman is matched to the man whose offer she was holding when the algorithm stopped (if any).

  That's why (final) acceptance was **deferred**

- Each man is matched to the woman he was temporarily matched when the algorithm stopped (if any).

Stability: existence

Theorem (David Gale & Lloyd Shapley, 1962)

The deferred acceptance algorithm always produces a stable matching.

D. Gale and L. Shapley (1962) "College admissions and the stability of marriage," *American Mathematical Monthly*, vol. 69, pp. 9–15.

Proof.

□

Stability: The algorithm were being used in practice

An excerpt of a letter from Gale to the NRMP.

> I'm sorry to trouble you with this request, but every time I talk on the subject someone always mentiones the National Matching Program and I feel it's time I found out what the relationship is in the instance between "theory and practice."

Stability: The algorithm were being used in practice

An excerpt of a letter from Elliott Peranson to D. Gale.

> You will note that the results indicated here match those that would be obtained from the "deferred - acceptance" algorithm outlined in your paper. However, I might point out that the NIRMP algorithm in fact uses the inverse procedure and produces the unique "college optimal" assignment rather than the "student optimal" assignment. This procedure more closely parallels the actual admissions process where a matching algorithm is not used.

Not only the theory could be applied in practice, it already had been more than 10 years earlier.

Will stable matchings occur in practice?

Empirical observation (natural experiment): for labor markets to survive, stability is a key property.

Market	Stable	Still in use
NRMP	yes	yes (new design 98-)
Edinburgh ('69)	yes	yes
Cardiff	yes	yes
Birmingham	no	no
Edinburgh ('67)	no	no
Newcastle	no	no
Sheffield	no	no
Cambridge	no	yes
London Hospital	no	yes
Medical Specialties	yes	yes (1/30 no)
Canadian Lawyers	yes	yes
Dental Residencies	yes	yes (2/7 no)
Osteopaths (-'94)	no	no
Osteopaths ('94-)	yes	yes
Reform rabbis	yes	yes
NYC highschool	yes	yes

Take-away

- A stable matching exists for each market: there is an algorithm to find one,
- the algorithm had already been in use 10 years earlier,
- stability is needed for organized markets to survive (evidence from natural experiment).

Deferred Acceptance algorithm: questions

- Is men-proposing equivalent to women proposing?
    - Straightforward answer.
- Which side is favored by the algorithm? Conflict or common interest?
    - Deep answer.

Deferred Acceptance algorithm: questions

The men-proposing is not equivalent to the women-proposing:

Example

P_{m_1}	P_{m_2}	P_{m_3}	P_{m_4}	P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}
w_1	w_2	w_3	w_4	m_4	m_3	m_2	m_1
w_2	w_1	w_4	w_3	m_3	m_4	m_1	m_2
w_3	w_4	w_1	w_2	m_2	m_1	m_4	m_3
w_4	w_3	w_2	w_1	m_1	m_2	m_3	m_4

- men proposing: $\mu_M = \begin{pmatrix} m_1 & m_2 & m_3 & m_4 \\ w_1 & w_2 & w_3 & w_4 \end{pmatrix}$
- women proposing $\mu_W = \begin{pmatrix} m_1 & m_2 & m_3 & m_4 \\ w_4 & w_3 & w_2 & w_1 \end{pmatrix}$

Deferred Acceptance algorithm: questions

- Coincidence of interest: every man prefers the men-proposing to the women-proposing,
- general phenomenon: every man (weakly) prefers the men-proposing to any other stable matching (optimality),
- general phenomenon: the common interest of the two sides are opposed on the set of stable matchings.

Definition

A stable matching is a men-optimal stable matching if every man (weakly) prefers it to any other stable matching.

Deferred Acceptance algorithm: optimality

Theorem (David Gale & Lloyd Shapley, 1962)

When all men and women have strict preferences, there always exist a men-optimal stable matching μ_M and a women-optimal stable matching μ_W. Furthermore, μ_M is produced by the deferred acceptance algorithm with men proposing. The matching μ_W is produced by the deferred acceptance algorithm with the women proposing.

μ_M is called the man-optimal matching.

Proof.

$\square$

Deferred Acceptance algorithm: optimality

- Matching obtained when running DA with men proposing:

  man-optimal matching.

- Matching obtained when running DA with women proposing:

  woman-optimal matching.

Stable matchings: opposition of interest

P_{m_1}	P_{m_2}	P_{m_3}	P_{m_4}	P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}
w_1	w_2	w_3	w_4	m_4	m_3	m_2	m_1
w_2	w_1	w_4	w_3	m_3	m_4	m_1	m_2
w_3	w_4	w_1	w_2	m_2	m_1	m_4	m_3
w_4	w_3	w_2	w_1	m_1	m_2	m_3	m_4

Is this an accident?

Stable matchings: opposition of interest

P_{m_1}	P_{m_2}	P_{m_3}	P_{m_4}	P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}
w_1	w_2	w_3	w_4	m_4	m_3	m_2	m_1
w_2	w_1	w_4	w_3	m_3	m_4	m_1	m_2
w_3	w_4	w_1	w_2	m_2	m_1	m_4	m_3
w_4	w_3	w_2	w_1	m_1	m_2	m_3	m_4

Is this an accident?

Stable matchings: opposition of interest

Proposition (Knuth)

When all men and all women have strict preferences,

all men (weakly) prefer the stable matching μ to the stable matching μ'

$$\Updownarrow$$

all women (weakly) prefer the stable matching μ' to the stable matching μ.

Proof.

$\square$

Stable matchings: opposition of interest

Corollary

When all men and women have strict preferences,

- *all men (weakly) prefer any stable matching to μ_W (men-pessimal stable matching)*
- *all women (weakly) prefer any stable matching to μ_M.*

Proof.

□

Stable matchings: opposition of interest

Warning!: there is no (always) common interest over all stable matchings, for each side.

P_{m_1}	P_{m_2}	P_{m_3}	P_{m_4}
w_1	w_2	$\boxed{w_3}$	$\boxed{w_4}$
$\boxed{w_2}$	$\boxed{w_1}$	w_4	w_3
w_3	w_4	w_1	w_2
w_4	w_3	w_2	w_1

P_{w_1}	P_{w_2}	P_{w_3}	P_{w_4}
m_4	m_3	m_2	m_1
m_3	m_4	m_1	m_2
$\boxed{m_2}$	$\boxed{m_1}$	m_4	m_3
m_1	m_2	$\boxed{m_3}$	$\boxed{m_4}$

Take-away

- There is a men(women)-optimal stable matching
- the common interest of the two sides of the market are opposed on the set of stable matchings.
- the women(men)-optimal stable matching is the men(women)-pessimal stable matching

Lattice of stable matchings: decomposition lemma

Definition

Given two stable matchings μ and μ', let M_μ denote the set of men who prefer μ to μ'. Analogously, define W_μ.

Lemma (Gale & Sotomayor, 1985)

When the preferences of all men and women are strict, the stable matchings μ and μ' are bijections between M_μ and $W_{\mu'}$.

Proof.

□

Lattice of stable matchings: decomposition lemma

Definition

Given two stable matchings μ and μ', let M_μ denote the set of men who prefer μ to μ'. Analogously, define W_μ.

Lemma (Gale & Sotomayor, 1985)

When the preferences of all men and women are strict, the stable matchings μ and μ' are bijections between M_μ and $W_{\mu'}$.

Proof.

□

Lattice of stable matchings

Theorem (Gale & Sotomayor, 1985)

When the preferences of all men and women are strict, the set of people who are single is the same for all stable matchings.

Proof.

□

Opposition of common interest over a break?

Lattice of stable matchings

Definition

Let μ and μ' be two matchings. Let $\mu \vee_M \mu'$ be defined on $M \cup W$ by:

- for each man m, $\mu \vee_M \mu'$ (m) is the more preferred mate of m between $\mu(m)$ and $\mu'(m)$,
- for each woman w, $\mu \vee_M \mu'$ (w) is the lees preferred mate of w between $\mu(w)$ and $\mu'(w)$.

In a precisely similar way, we define $\mu \wedge_M \mu'$ which gives each man his less preferred mate and each woman her more preferred mate.

Lattice of stable matchings

Lemma (Lattice theorem (Convey))

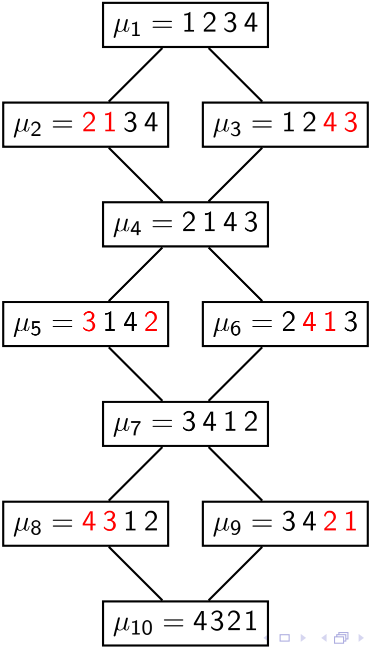
When all men and women have strict preferences and μ and μ' are stable matchings, $\lambda = \mu \vee_M \mu'$ and $\mu \wedge_M \mu'$ are matchings and they are stable.

Proof.

$\square$

Corollary

When all men and all women have strict preferences, the set of stable matchings forms a lattice w.r.t the common preferences of men or women.


$\mu_1 = 1\,2\,3\,4$
$\mu_2 = 2\,1\,3\,4$
$\mu_3 = 1\,2\,4\,3$
$\mu_4 = 2\,1\,4\,3$
$\mu_5 = 3\,1\,4\,2$
$\mu_6 = 2\,4\,1\,3$
$\mu_7 = 3\,4\,1\,2$
$\mu_8 = 4\,3\,1\,2$
$\mu_9 = 3\,4\,2\,1$
$\mu_{10} = 4321$

When preferences need not be strict

Not hold	Hold
Existence of optimal stable matchings Opposition of common interest Decomposition lemma Lattice structure	Existence of stable matchings

Centralized markets: incentive

Consider a market organized by a matchmaker (computer service):

- National Resident Matching Program
- Public school programs,
- etc.

The matchmaker collects preferences and arranges matches. But preferences are private information.

Is it in the best interest of each agent to state his or her preferences to the matchmaker?

Centralized markets: incentive

Consider a market organized by a matchmaker (computer service):

- National Resident Matching Program
- Public school programs,
- etc.

The matchmaker collects preferences and arranges matches. But preferences are private information.

Is it in the best interest of each agent to state his or her preferences to the matchmaker?

Centralized market: incentive

Definition

- A strategy is a dominant strategy for an agent in the mechanism φ if it is a best response to any strategy of the other agents,
- A mechanism is strategy-proof if submitting his true preferences is a dominant strategy for each agent.

Model

We consider the following mechanism:

1. Men and woman submit (simultaneously) their preferences,
2. a mechanism (or algorithm) uses the submitted preferences,
3. the matching is announced.

$$\boxed{\text{Preferences } P} \longrightarrow \boxed{\text{Mechanism } \varphi} \longrightarrow \boxed{\text{Matching } \mu}$$

Question: If the mechanism chooses a stable matching according to the submitted preferences, do men and women have the incentive to submit their true preferences?

Model

We consider the following mechanism:

1. Men and woman submit (simultaneously) their preferences,
2. a mechanism (or algorithm) uses the submitted preferences,
3. the matching is announced.

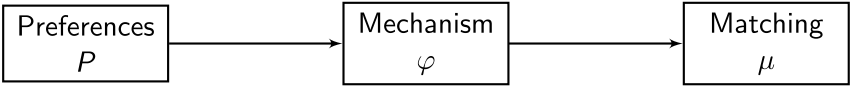
$$\boxed{\text{Preferences } P} \longrightarrow \boxed{\text{Mechanism } \varphi} \longrightarrow \boxed{\text{Matching } \mu}$$

Question: If the mechanism chooses a stable matching according to the submitted preferences, do men and women have the incentive to submit their true preferences?

Model

We consider the following mechanism:

1. Men and woman submit (simultaneously) their preferences,
2. a mechanism (or algorithm) uses the submitted preferences,
3. the matching is announced.



Question: If the mechanism chooses a stable matching according to the submitted preferences, do men and women have the incentive to submit their true preferences?

Incentives: impossibility

Theorem (Roth, 1982)

No stable matching mechanism exists for which stating the true preferences is a dominant strategy for every agent.

Proof by example.

P_{m_1}	P_{m_2}
w_1	w_2
w_2	w_1

P_{w_1}	P_{w_2}
m_2	m_1
m_1	m_2

$$\mu_M(m_1) = w_1 \quad \text{and} \quad \mu_M(m_2) = w_2$$

$$\mu_W(m_1) = w_2 \quad \text{and} \quad \mu_W(m_2) = w_1$$

$\square$

Incentives: impossibilities

When do agents have an incentive to misrepresent their preferences to any stable matching mechanism?

Which agents have this incentive?

Incentives: impossibilities

Theorem

When any stable mechanism is applied to a marriage market in which preferences are strict and there is more than one stable matching, then (assuming that others tell the truth)

every agent can misrepresent his preferences in such a way to be matched to his most preferred achievable mate under the true preferences.

Proof.

□

Does an agent who received his most preferred achievable mate have the incentive to misrepresent his or her preferences?

Incentives: possibility

Incentives facing the men when the men-optimal stable matching mechanism is employed.

Theorem (Dubins and Freedman, 1981)

Let P be the truth preferences of the agents, and P' differ from P in that some coalition M' of men misrepresented their preferences.

There is no matching μ, stable under P', which every member of M' prefers to μ_M.

Corollary

The men-optimal stable matching mechanism is strategy-proof for men.

Incentives: possibility

Incentives facing the men when the men-optimal stable matching mechanism is employed.

Theorem (Dubins and Freedman, 1981)

Let P be the truth preferences of the agents, and P' differ from P in that some coalition M' of men misrepresented their preferences.

There is no matching μ, stable under P', which every member of M' prefers to μ_M.

Corollary

The men-optimal stable matching mechanism is strategy-proof for men.

Incentives: possibility

Lemma (Blocking lemma)

Let μ be an individually rational matching with respect to strict preferences P, and let M' be the set of all men who prefer μ to μ_M.

If M' is nonempty, there is a pair (m, w) that blocks μ such that $m \notin M'$ and $w \in \mu(M')$.

Proof of the lemma.

$\square$

Corollary (Weak Pareto optimality for men)

There is no individually rational matching μ, stable or not, that each man prefers to μ_M

Incentives: possibility

Lemma (Blocking lemma)

Let μ be an individually rational matching with respect to strict preferences P, and let M' be the set of all men who prefer μ to μ_M.

If M' is nonempty, there is a pair (m, w) that blocks μ such that
$$m \notin M' \text{ and } w \in \mu(M').$$

Proof of the lemma.

$\square$

Corollary (Weak Pareto optimality for men)

There is no individually rational matching μ, stable or not, that each man prefers to μ_M

Take-away

- the Deferred Acceptance algorithm (DA) produces a stable matching:
  - the most preferred stable matching for the proposing side,
  - the least preferred stable matching for the receiving side,
- DA is strategyproof for the proposing side, but not for the receiving side,
- we cannot have, in general, strategyproofness for both sides and stability.
- any man or woman who did not receive his side optimal-stable's mate at a stable matching mechanism can misrepresent his preferences in such a way to receive it.

Next lecture

If you are not happy for today's lecture, I hope you will be compensated next time:

- Housing market & House allocation
  - new market design problems