CSC304: Algorithmic Game Theory and Mechanism Design
Fall 2016

Allan Borodin (instructor)
Tyrone Strangway and Young Wu (TAs)

November 16, 2016

Lecture 18

- Announcements
  - I have posted the first 4 questions for the last assignment.
  - I should note that students are responsible for material disucssed in the lectures and tutorials whether or not that material is available in the lecture slides or text.
  - Announcement about Career Mentorship on next slide
- Todays agenda
  - Start mechanism design without money; stable matching

Career mentorship program

Computer Science Career Mentorship Program 2016/17

Our **Career Mentorship Program** connects undergraduates (Major or Specialist) with supportive alumni of computer science, who are working in the field, and want to share their academic and career experiences with a student mentee who is preparing for their own path.

The program runs from **January to April**. Students are also required to attend a prep session in December prior to the first event in January, and submit reflective exercises throughout the program. Undergraduates who successfully complete all requirements will receive Co-Curricular Record (CCR) credit.

Apply by November 18, 2016

alumni.artsci.utoronto.ca/dcs-mentee-application

Computer Science
UNIVERSITY OF TORONTO

Mechanism design without money

We have now discussed parts of chapters 14-17 dealing with various topics concerning different types of auctions and mechanisms for such auctions. These topics naturally involve the transfer of money.

The remaining (approximately one third) part of the course falls under the heading of "mechanism design without money". That is, mechanisms will try to achieve desireable outcomes but without any transfer of money. As before, self-interested agents have private information. Agents may have preferences (rather than valuations) as their private information.

In this regard we will discuss three topics in this order:

1. Stable matching: the stable marriage problem (chapter 10)
2. Social choice: Voting (chapter 13)
3. Fair division (chapter 11)

One could say that stable matching is also part of "social choice" but usually one has a more restrictive meaning for this term. But beyond voting, it is reasonable to say that topics like recommendation systems and peer evaluation are part of social choice.

The stable marriage problem

The stable marriage problem originates with a 1960 paper by Gale and Shapley. Lloyd Shapley was the co-winner (with David Roth) of the 2012 Nobel prize in Economics for "the theory of stable allocations and the practice of market design".

We will begin with the classic statement of the problem (from the 1960 paper) and then discuss some of the more practical issues that Roth and others have addressed. This remains a topic of current interest both theoretically and practically.

The stable marriage problem

The stable marriage problem originates with a 1960 paper by Gale and Shapley. Lloyd Shapley was the co-winner (with David Roth) of the 2012 Nobel prize in Economics for "the theory of stable allocations and the practice of market design".

We will begin with the classic statement of the problem (from the 1960 paper) and then discuss some of the more practical issues that Roth and others have addressed. This remains a topic of current interest both theoretically and practically.

Informally, in the classic problem, the goal of the mechanism is to match men and women in such a way that matched couples have no reason to abandon their match.

Beyond any possible use of the classic male-female matching application, there are many other important applications. One very practical application for stable matching is the NRMP (National Residents Matching Program) that matches residents with hospital programs.

The two-sided marriage/matching problem

We are given two sets of agents X and Y. (In some problems, we could have $X = Y$, but our focus is when X and Y are disjoint.)

Every $x \in X$ has some preference ordering over Y, and every $y \in Y$ has some preference ordering over X.

Like matching markets, the goal is to match agents in X with agents in Y to satisfy some objective. *But unlike the one-sided problem* now both X and Y are strategic.

And unlike matching markets, agents do not have valuations but rather have a partial or full ranking of the agents "on the other side".

The two-sided marriage/matching problem

We are given two sets of agents X and Y. (In some problems, we could have $X = Y$, but our focus is when X and Y are disjoint.)

Every $x \in X$ has some preference ordering over Y, and every $y \in Y$ has some preference ordering over X.

Like matching markets, the goal is to match agents in X with agents in Y to satisfy some objective. *But unlike the one-sided problem* now both X and Y are strategic.

And unlike matching markets, agents do not have valuations but rather have a partial or full ranking of the agents "on the other side".

Many applications. Perhaps the most studied application is the US NRMP. X is the set of medical residents, $Y =$ hospital residency programs. In the US NRMP, ≈ 35000 residents, ≈ 4000 programs. Each of these programs has many positions.

Canada and Scotland also use residency matching programs.

Some other applications

- X is a set of companies and Y is a set of applicants Y.
- X is a set of gradaute students and Y is a set of faculty supervisors. (Apparently this is done at Duke University.)
- New York City and Boston both do centralized school matching.
- Students needing to share dorm rooms. Here is an application where $X = Y$.
- Assigning papers to reviewers.

The classic marriage problem

In the classic Gale and Shapley model, there are say n males and n' females. For simplicity, we can assume that $n = n'$ but this is not at all essential. Each man has a total ranking with respect to the set of women, and each woman has a total ranking over the set of men. The goal is to find a "stable" perfect (or maximum size when $n \neq m$) matching.

Note that unlike Donald Trump, neither the men nor the women give numeric values but rather each has a total (and private) ranking for possible mates. This is admittedly an unrealistic assumption as usually, individuals would only have a partial ranking. We will later briefly discuss this and other complicating but realistic issues.

The classic problem also assumes that the goal is a 1-1 matching whereas most applications are many to one problems. For example, a company may want to hire many employees or a hospital may have many residency positions.

Preference orderings

A total preference ordering is just a ranked list of all potential partners. Staying with the classical example, we will let $\succ_m$ denote the preference ordering (over the set W of women) of a male $m \in M$ and similarly $\succ_f$ will denote the preference ordering of a female $f \in W$ over the set M of men.

Preference orderings

A total preference ordering is just a ranked list of all potential partners. Staying with the classical example, we will let $\succ_m$ denote the preference ordering (over the set W of women) of a male $m \in M$ and similarly $\succ_f$ will denote the preference ordering of a female $f \in W$ over the set M of men.

Note that for now we are assuming that the preference ordering is total; that is, for every u, v either $u \succ v$ or $v \succ u$. (This is the exclusive "or".) And for now we are not allowing u and v to be equally preferred.

A total ranking for males and females

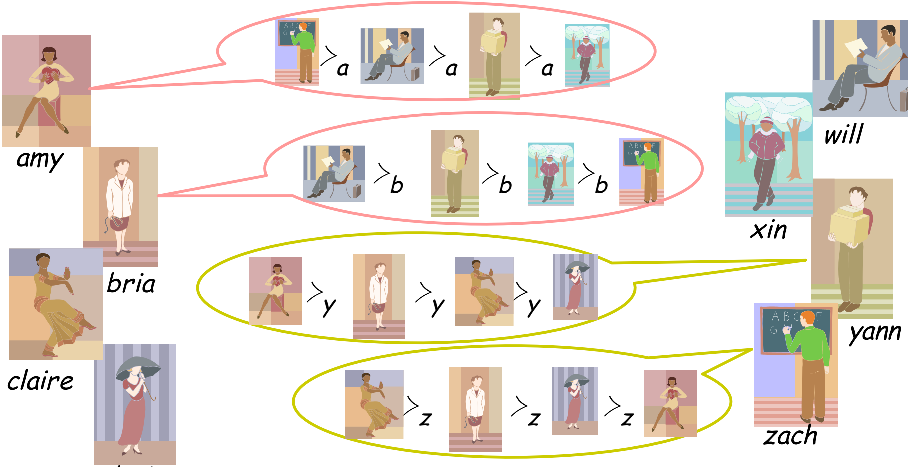


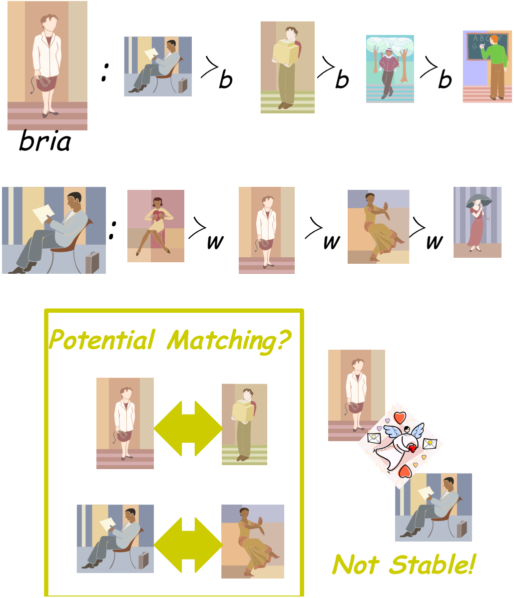
Figure: Everyone has a total ranking for the opposite gender

Stable matchings and blocking pairs

Our goal is to find a perfect (or maximum) matching that has some desireable properties. What is desireable?

- As stated there are no valuations so social welfare is not well defined in this context. If each agent had a distinct value for every choice (as in matching markets), then that induces a total ordering. But it is hard enough to assume preferences much less complete valuations. Conversely, we could give a somewhat ad-hoc valuation to the matched pairs (m, w) by providing a scoring rule for ones position in the ranking. But generally speaking, we do not consider social welfare in this problem.
- On the other hand, Pareto optimality is well defined in this context.
- Our main criteria in this context is "stability" which informally means that no pair of individuals would rather be with each other than with their assigned match. Such a pair is called a *blocking pair* in that the pair is blocking stability as they have reason and the ability to abandon the matching.

Example of a blocking pair



Defining stability and blocking pairs

To simplify notation, we will abuse notation and let $\mu : M \to W$ and $\mu : W \to M$ denote the matching between men and women. For a matching we require $\mu(m) = w$ iff $\mu(w) = m$.

Defining stability and blocking pairs

To simplify notation, we will abuse notation and let $\mu : M \to W$ and $\mu : W \to M$ denote the matching between men and women. For a matching we require $\mu(m) = w$ iff $\mu(w) = m$.

A pair (m, w) blocks the matching μ if $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. In words, m and w would rather be with each other than with their currently assigned partners.

Defining stability and blocking pairs

To simplify notation, we will abuse notation and let $\mu : M \to W$ and $\mu : W \to M$ denote the matching between men and women. For a matching we require $\mu(m) = w$ iff $\mu(w) = m$.

A pair (m, w) blocks the matching μ if $w \succ_m \mu(m)$ and $m \succ_w \mu(w)$. In words, m and w would rather be with each other than with their currently assigned partners.

A matching μ is stable if it is not blocked by any pair (m, w). That is, if $w \succ_m \mu(m)$, then $m \prec_w \mu(w)$; and similarly, if $m \succ_w \mu(w)$ then $w \prec_m \mu(m)$. In words, while some man (resp. woman) might prefer another woman (resp. man), the feeling is not mutual.

Some examples of stable and unstable matchings

Man	1st	2nd	3rd
x	a	b	c
y	b	a	c
z	a	b	c

Woman	1st	2nd	3rd
a	y	x	z
b	x	y	z
c	x	y	z

Match 1
a-x
b-y
c-z

It is stable (x,y don't want to move; a, b won't move except to x,y; so c,z stuck)

Match 2
a-y
b-x
c-z

It is stable (a,b don't want to move; x, y won't move except to a,b; so c,z stuck)

Match 3
a-z
b-y
c-x

Unstable (b,x form a blocking pair)

(a,x also form a blocking pair)

Figure: Two stable and one unstable example

Stability as a type of equilibrium

As agents are self interested, they are always looking to improve their "payoff" (i.e. the ranking of a partner). A stable matching means that no one is able to improve their payoff. However, this is different than Nash equilbria, in that it "takes two to tango". No one can improve their situation on their own. One can call this "coalitional stability".

Stability as a type of equilibrium

As agents are self interested, they are always looking to improve their "payoff" (i.e. the ranking of a partner). A stable matching means that no one is able to improve their payoff. However, this is different than Nash equilbria, in that it "takes two to tango". No one can improve their situation on their own. One can call this "coalitional stability".

Important "aside": In our setting, no one is unacceptable. That is, we are assuming what is, in many cases, the unrealistic situation that even if you get your worst choice, you would still rather be with someone than not get matched. (For some applications this might not be so unreasonable. For example, maybe getting into some University is better than not going at all?)

If we want to accomodate the fact that individuals may have a threshold for acceptance, then (for example) we can extend the range of a preference $\succ_m$ to $W \cup \{m\}$ where say $w_1 \succ_m w_2 \succ_m w_3 \succ_m m \succ_m w_4 \succ_m w_5$ indicates that m would rather be unmatched than be with w_4 or w_5.

The Gale-Shapley Deferred Acceptance algorithm

It turns out that there is a relatively efficient was to construct a perfect (or maximum size) matching that is stable.

The Deferred Acceptance (DA) algorithm comes in two versions; namely, the Female Proposes (FPDA) and Male Proposes (MPDA) algorithms.

For simplicity, lets assume $n = n'$ so that our goal is to establish a perfect matching that is stable. (If $n \neq n'$, then the goal is to find a stable matching of size $\min(n, n')$.)

For definiteness, lets consider the FPDA. The MPDA is the same with the roles of men and women reversed.

The Female Proposes Delayed Acceptance FPDA algorithm

The FPDA algorithm

Initialize the *currently engaged CE list* to be empty;
that is, all males and females are *not engaged*.
Initialize every woman w's *already proposed* AP_w *list* to be empty
In rounds DO until every female is engaged

1. Every not engaged female w proposes to the man $m \notin AP_w$ that is most preferred in her preference ranking $\succ_w$.
2. For each man m, let $PT_m(t)$ be the set of women who have proposed to m in this round t plus the woman w he is currently engaged to (if engaged)
   - 2A. Let w^* be the woman most preferred in $P_m(t)$
   - 2B. Add (m, w^*) to CE removing (m, w) from CE if $w \neq w^*$ and (m, w) were engaged at the start of the round

End DO

FPDA example 1: round 1

Wmn	1st	2nd	3rd	4th
a	x	y	z	w
b	y	x	w	x
c	x	z	y	w
d	y	w	x	z

A * means: already proposed to that man

Man	1st	2nd	3rd	4th
w	d	b	a	c
x	b	a	d	c
y	b	a	c	d
z	d	b	c	a

Round 1

Proposals:	New Engagements:
a: x	w: −
b: y	x: a
c: x	y: b
d: y	z: −

FPDA example 1: round 2

Wmn	1st	2nd	3rd	4th
a	x*	y	z	w
b	y*	x	w	x
c	x*	z	y	w
d	y*	w	x	z

A * means: already proposed to that man

Man	1st	2nd	3rd	4th
w	d	b	a	c
x	b	a	d	c
y	b	a	c	d
z	d	b	c	a

Round 2

Current:	Proposals:	New Engagements:
w: –	a: –	w: d
x: a	b: –	x: a
y: b	c: z	y: b
z: –	d: w	z: c

Done - Stable:

a:x
b:y
c:z
d:w

FPDA example 2: round 1

Wmn	1st	2nd	3rd	4th
a	x	y	z	w
b	y	x	w	x
c	x	**y**	**z**	w
d	y	w	x	z

A * means: already proposed to that man

Man	1st	2nd	3rd	4th
w	d	b	a	c
x	b	a	d	c
y	**c**	**b**	**a**	**d**
z	d	b	c	a

Round 1

Proposals:	New Engagements:
a: x	w: –
b: y	x: a
c: x	y: b
d: y	z: –

FPDA example 2: round 2

Wmn	1st	2nd	3rd	4th
a	x*	y	z	w
b	y*	x	w	x
c	x*	y	z	w
d	y*	w	x	z

A * means: already proposed to that man

Man	1st	2nd	3rd	4th
w	d	b	a	c
x	b	a	d	c
y	c	b	a	d
z	d	b	c	a

Round 2

Current:	Proposals:	New Engagements:
w: –	a: –	w: d
x: a	b: –	x: a
y: b	c: y	y: ~~b~~ c
z: –	d: w	z: –

b is “jilted”

FPDA example 2: round 3

Wmn	1st	2nd	3rd	4th
a	x*	y	z	w
b	y*	x	w	x
c	x*	y*	z	w
d	y*	w*	x	z

A * means: already proposed to that man

Man	1st	2nd	3rd	4th
w	d	b	a	c
x	b	a	d	c
y	c	b	a	d
z	d	b	c	a

Round 3

Current:	Proposals:	New Engagements:
w: d	a: –	w: d
x: a	b: x	x: ~~a~~ b
y: ~~b~~ c	c: –	y: ~~b~~ c
z: –	d: –	z: –

a is "jilted"

FPDA example 2: round 4

Wmn	1st	2nd	3rd	4th
a	x*	y	z	w
b	y*	x*	w	x
c	x*	y*	z	w
d	y*	w*	x	z

A * means: already proposed to that man

Man	1st	2nd	3rd	4th
w	d	b	a	c
x	b	a	d	c
y	c	b	a	d
z	d	b	c	a

Round 4

Current:	Proposals:	New Engagements:
w: d	a: y	w: d
x: ~~a~~ b	b: –	x: ~~a~~ b
y: ~~b~~ c	c: –	y: ~~b~~ c
z: –	d: –	z: –

a's proposal not accepted by y
(no change)

FPDA example 2: round 5

Wmn	1st	2nd	3rd	4th
a	x*	y*	z	w
b	y*	x*	w	x
c	x*	y*	z	w
d	y*	w*	x	z

A * means: already proposed to that man

Man	1st	2nd	3rd	4th
w	d	b	a	c
x	b	a	d	c
y	c	b	a	d
z	d	b	c	a

Round 5

Current:	Proposals:	New Engagements:
w: d	a: z	w: d
x: ~~a~~ b	b: –	x: ~~a~~ b
y: ~~b~~ c	c: –	y: ~~b~~ c
z: –	d: –	z: a

Stable:
a:z
b:x
c:y
d:w

Why does FPDA work?

We need to show that the FPDA always terminates and ends in a stable matching. Both termination and stability follow from a few observations:

- The sequence of women to which any man is engaged is strictly improving This is obvious since men are only accepting better and better offers
- The sequence of men proposed to by any woman gets worse in each new proposal This is again obvious since a woman who has to propose again must go to the next best person in her list.
- These observations not only guarantee termination but show that at most n^2 proposals are needed. Why? There are instances where the number of proposals come pretty close to this bound.
- The claim is that there cannot be a blocking pair (m, w) for the match μ produced by the FPDA algorithm. Why?

Why we cannot have a blocking pair

Suppose (m, w) is a blocking pair. Then w prefers m to $\mu(w)$. This means she would have proposed to m before proposing to $\mu(w)$. But then what would have happened if m prefers w to $\mu(m)$?

Why we cannot have a blocking pair

Suppose (m, w) is a blocking pair. Then w prefers m to $\mu(w)$. This means she would have proposed to m before proposing to $\mu(w)$.
But then what would have happened if m prefers w to $\mu(m)$?

Case 1: If w proposed before $\mu(m)$, then m would not have accepted $\mu(m)$'s proposal since he would be engaged to either w or someone even more preferred than w.

Case 2: If w proposed after $\mu(m)$, then he would have rejected/jilted $\mu(m)$.

This then shows that μ could not have been the matching so we conclude there is no blocking pair.