About this class

Two-Sided Matching (mostly from Roth and Sotomayor)

Basic Structure

Two separate sides of the market – hospitals and medical students, men and women, colleges and students

Agents on both sides of the market have preferences over those on the other side

For the most part we'll be thinking about non-transferable utility

Unraveling

NRMP: Hospitals started making offers earlier and earlier and asking for earlier and earlier commitments from potential residents

This makes sense for all parties involved given the behavior of the other parties, but everyone would have preferred a different setup

Need to analyze this problem in terms of incentives and mechanisms

Rationality and Stability

Let's simplify by thinking about one-to-one matchings, say between men and women

Matching is individually rational for each agent if they'd rather be matched with the person they are matched with than be single

Consider the following setup – 3 men and 3 women with preferences as follows:

$$
\begin{array}{ll}
m_1 : w_2 > w_1 > w_3 & w_1 : m_1 > m_3 > m_2 \\
m_2 : w_1 > w_3 > w_2 & w_2 : m_3 > m_1 > m_2 \\
m_3 : w_1 > w_2 > w_3 & w_3 : m_1 > m_3 > m_2
\end{array}
$$

Everyone prefers to be matched with someone rather than single

Consider the following matching:

$$\begin{array}{ccc} w_1 & w_2 & w_3 \\ m_1 & m_2 & m_3 \end{array}$$

It's individually rational, but m_1 and w_2 would rather be with each other...

This leads to the concept of stability: there is no pair of agents who would rather be with each other than with those they are matched with

Stability: Roommates

Four students need to get split into two rooms of two people each. No one likes one of the students, d, whose preferences are arbitrary:

$$\begin{aligned} a &: \; b > c > d \\ b &: \; c > a > d \\ c &: \; a > b > d \\ d &: \end{aligned}$$

Suppose a and b are matched in one room. b would rather be with c, and c would rather be with b than with d...you can see that this sort of thing holds for all matches – this means there are no stable matches!

This won't happen in the two-sided version...

The Gale-Shapley Algorithm

Here's the algorithm, men proposing version:

While there is a free man m who hasn't proposed to every woman who is acceptable to him:

- Let w be the highest ranked woman in m's list to whom he hasn't proposed. m proposes to w. If w is free m and w become engaged. Else, if w prefers m to her current fiancee m' then m and w become engaged and m' becomes free. Else m remains free

Proof that the matching is stable:

Suppose m and w are not married but m prefers w to his own mate. Then he must have proposed to her before proposing to his current mate. Then she must have been either rejected by her then (indicating she prefers someone else) or dumped by her later (again indicating that she prefers someone else). Then by transitivity of preferences, she must prefer her current mate.

Women- and Men-Optimal Matchings

$$
\begin{array}{ll}
m_1 : w_1 > w_2 > w_3 & w_1 : m_1 > m_2 > m_3 \\
m_2 : w_1 > w_2 > w_3 & w_2 : m_1 > m_3 > m_2 \\
m_3 : w_1 > w_3 > w_2 & w_3 : m_1 > m_2 > m_3
\end{array}
$$

It's worth thinking about a couple of things. First, all the men would like to be with w_1 – there's some incentive for them to stay outside the stability system, but, hey, what can you do?

If we restrict our attention to stable matchings, there's still going to be some interesting stuff. Consider two possible stable matchings:

First,

$$
\begin{array}{ccc}
w_1 & w_2 & w_3 \\
m_1 & m_2 & m_3
\end{array}
$$

Second,

$$\begin{array}{ccc} w_1 & w_3 & w_2 \\ m_1 & m_2 & m_3 \end{array}$$

These are both stable, m_1 and w_1 are indifferent among the two matchings. m_2 and m_3 both prefer the first one, and w_2 and w_3 both prefer the second one. Interestingly, the first one is generated by the men-proposing mechanism, and the second one is generated by the women-proposing mechanism!

A men-optimal stable matching is one that is at least as good as any other stable matching for every man, and a women-optimal matching is defined similarly.

Now let's show that the men-proposing mechanism achieves the men-optimal stable matching. Let a woman be *achievable* for a man if they are paired in some stable matching. We'll show that no man is ever rejected by an achievable woman with men proposing.

By induction: suppose man m has not yet been rejected by an achievable woman and then woman w rejects him. If she's not yet paired with someone that means he's unacceptable to her and therefore she's not achievable. If not, she rejected him in favor of some m'.

Then it must be that m' prefers w to any woman except those who have already rejected him (by the structure of the algorithm). Then there is no woman who is both achievable for m' and ranked higher on his list than w (by the inductive assumption).

Suppose there were some matching that paired w and m and everyone else to an achievable mate, then w would prefer m' and m' must prefer w (because she's the highest achievable for him), and thus the matching must be unstable. This shows that w is not achievable for m.

Obviously the exact same thing holds for women in the women-proposing mechanism.