Two-Sided Matching

Hanzhe Zhang

Tuesday, November 19, 2013

We will study the two-sided matching problems today, and focus on the strand of research that has provided theoretical foundations to the field of market design combining engineering, economics, and experiments to solve real world allocation problems. Successful applications include National Residence Medical Program (NRMP) for assigning new medical doctors to their first positions in hospitals, kidney exchanges, and middle and high school choices. These theoretical and practical contributions have profound impacts and are well recognized. Lloyd Shapley and Alvin Roth were awarded the Nobel Prize in 2012 for "the theory of stable allocations and the practice of market design"; David Gale should have been the third recipient if he were alive.

I will introduce the concept of stability and the deferred acceptance algorithm proposed by David Gale and Lloyd Shapley in 1962 that finds a stable outcome in two-sided matching situations, and discuss the game-theoretic issues of the algorithm studied by Alvin Roth and others.

For the first half of the lecture, no knowledge of game theory is needed for understanding stability and the algorithm; but for the second half, I will turn the algorithm into part of a noncooperative game when the agents are required to report their preferences as inputs of the algorithm.

1 Setup

Examples of two-sided matching problems and markets are pervasive. Men and women choose among a pool of candidates to date and marry. Workers choose to work for different firms who hire different types of workers according to their needs. Colleges decide what students to attract and accept and students apply to different colleges. Course scheduling is another example: each student wants to take three or four classes, and each class accommodates a fixed number of people.

There are many common features among these markets; especially important are three. First, they are two-sided; there are two disjoint sets of agents: men and women, workers and firms, students and colleges, and students and classes. Second, everyone has a constraint (quota) on the number of partners. Men can have only one wife and every woman marries only one husband.

Students can only attend one college, and each college has a physical or administrative capacity on how many students to take. Third, the agents usually have heterogeneous preferences over their possible mates. Men and women have different criteria for their marriage partners. Students may like different colleges, and colleges may want different student compositions. CalTech wants science oriented students, but West Point may want those who have demonstrated leadership skills.

These similarities lead us to a general setup of these matching markets. First, let M and W denote the two disjoint sets of agents on each side of the market. Let $m \in M$ and $w \in W$ denote a generic agent from each side of the market. Second, each agent has an individualized quota q on the maximum number of partners one can have. If everyone has quota 1, then we have a one-to-one matching market, like the marriage market. If everyone on one side of the market has quota 1 and everyone on the other side has bigger quota, then it is a many-to-one matching market; for example, the college admissions process and the labor market. Many-to-many matching markets, like the course scheduling, can be similarly defined. In the following presentation, we focus on one-to-one matching markets. Finally, each person has a preference relation P on the agents on the other side of the market and the situation being unmatched. For simplicity, we let the preference to be strict. In the first part of the lecture, we take these preferences as given.

Let's take an example of two boys, Alex and Bob, and two girls, Alice and Betty, pairing up to attend a (traditional) prom. They can only go with an opposite gender, and at most take one partner. Alex and Bob constitute the set M and Alice and Betty the set W. Their preferences are defined as follows. Both Alex and Bob would like to go with Alice than with Betty. Alice wants to go with Alex more than with Bob and Betty wants to go with Bob than with Alex. And all of them prefer to go with someone than alone. The preferences P can be written as follows.

$$\text{Alice } P_{\text{Alex}} \text{ Betty } P_{\text{Alex}} \emptyset$$

$$\text{Alice } P_{\text{Bob}} \text{ Betty } P_{\text{Bob}} \emptyset$$

$$\text{Alex } P_{\text{Alice}} \text{ Bob } P_{\text{Alice}} \emptyset$$

$$\text{Bob } P_{\text{Betty}} \text{ Alex } P_{\text{Betty}} \emptyset$$

For example, (Alice $\leftrightarrow$ Alex, Bob $\leftrightarrow$ Betty) is a match, and (Alice $\leftrightarrow$ Bob, Betty $\leftrightarrow \emptyset$, Alex $\leftrightarrow \emptyset$) is a match, but (Alice $\leftrightarrow$ Betty, Bob $\leftrightarrow$ Alex) is not a match, because same sexes cannot go together, and (Alice $\leftrightarrow$ (Alex,Bob), Betty $\leftrightarrow \emptyset$) is not feasible for Alice exceeds her quota.

Let's consider another example with four students α, β, γ and δ, and two colleges A and B. Each student can attend one college, and suppose A can take two students and B and only take 1. Suppose β, γ and δ all prefer A to B, but α prefers B to A. College A ranks the students alphabetically: $\alpha P_A \beta P_A \gamma P_A \emptyset P_A \delta$, and college B ranks almost reverse-alphabetically, $\gamma P_B \beta P_B \alpha P_B \emptyset P_B \delta$

where $\emptyset \succ_B \delta$ means that B finds δ to be unacceptable. (α, β) attending A and γ attending B is possible feasible college assignment, but $(\alpha, \beta, \gamma, \delta)$ all attending college B is not feasible because B can only take one student.

A matching outcome μ is formally defined as follows.

Definition 1. A (feasible) **matching** μ is a correspondence[1], $\mu : M \cup W \to 2^{M \cup W}$ such that for all $m \in M$ and $w \in W$,

- $\mu(m) \subseteq W$ and $|\mu(m)| \leq q_m$, $\mu(w) \subseteq M$ and $|\mu(w)| \leq q_w$, and
- $m \in \mu(w)$ if and only if $w \in \mu(m)$.

2 Stability

After developing the setup and defining the match, we want to use it to study the aforementioned markets. First of all, what is a reasonable and sustainable outcome in any market? Let's return to our prom example. Only a few outcomes are available. If there are two pairs then the only outcomes are (Alex $\leftrightarrow$ Alice, Bob $\leftrightarrow$ Betty) and (Alex $\leftrightarrow$ Betty, Bob $\leftrightarrow$ Alice). If there is one pair, there are a few possible matches: (Alex $\leftrightarrow \emptyset$, Alice $\leftrightarrow \emptyset$, Bob $\leftrightarrow$ Betty) etc. Or everyone goes alone: (Alex $\leftrightarrow \emptyset$, Alice $\leftrightarrow \emptyset$, Bob $\leftrightarrow \emptyset$, Betty $\leftrightarrow \emptyset$). Since no one wants to go alone, having everyone going alone does not seem to be a good and sustainable outcome. In fact, having any of them going alone is undesirable. That leaves us with two possible matches: (Alex $\leftrightarrow$ Alice, Bob $\leftrightarrow$ Betty) and (Alex $\leftrightarrow$ Betty, Bob $\leftrightarrow$ Alice). Alex and Alice are each other's favorite partners, so it is a more reasonable and desirable outcome than the other pairing. In the other pairing in which Alice goes with Bob and Alex goes with Betty, Alex and Alice would rather like to go with each other and they have no reason not to do so. We formalize the blocking pair notion.

Definition 2 (Blocking Pair)**.**

- Agents $m \in M$ and $w \in W$ **block** a one-to-one matching μ if $m P_w \mu(w)$ and $w P_m \mu(m)$.
- $m \in M$ and $w \in W$ **block** a many-to-many matching μ if $m \notin \mu(w)$ and $w \notin \mu(m)$, and
  - If $|\mu(m)| = q_m$, $w P_m w'$ for some $w' \in \mu(m)$, and if $|\mu(m)| < q_m$, $w P_m \emptyset$
  - If $|\mu(w)| = q_w$, $m P_w m'$ for some $m' \in \mu(w)$, and if $|\mu(w)| < q_w$, $m P_w \emptyset$
- μ is **unblocked** if there does not exist a pair of agents that block μ.

[1] 2^X denotes the power set of X, so an element of 2^X is a subset of X, possibly the empty set.

Furthermore, a match should be **individually rational**, i.e. no one should be matched with an unacceptable partner. In the college example, both A and B do not find δ acceptable. If δ is matched to A, A can **block** the match by kicking δ out.

Definition 3. μ is **individually rational** if for all $m \in M$, $w \in W$, $w' R_m \emptyset$ for all $w' \in \mu(m)$ and $m' R_w \emptyset$ for all $m' \in \mu(w)$.

We call a matching **stable** if the matching outcome is blocked by no agent alone and no pair of agents, and **unstable** otherwise.

Definition 4. A matching μ is **stable** if it is not blocked by any individual or any pair of agents. A matching is **unstable** if it can be blocked by an individual or a pair of individual.

Why is stability a reasonable outcome, and an outcome we want to achieve? First, by definition, no one can improve upon a stable outcome. In the marriage market, it means no divorce happens under a stable outcome. Return to our prom example. When Alex and Alice go together, even though Bob still likes to go with Alice, Alice does not want to switch partners. However, if Bob and Alice plan to pair, Alex can come steal Alice away; in fact, Alice has the incentive to ditch Bob by herself.

So far, it is still a thought experiment that generates some nice properties of stable outcomes and you may still hesitate to accept its importance. Let me give you a real world example - a brief history of the National Residence Matching Program (NRMP). In the 1950s, medical doctor students in their final year of study find jobs on their own and hospitals advertised openings by themselves. The competition was very fierce. It was so fierce that they signed contracts very early - before doctors had even any on-hand experience with actual medical practices. Well, it turned out that there was a low correlation between being a good student and being a good doctor. Many students fainted when they saw blood, for extreme examples. Inefficiencies were huge and growing, and it was a significant social waste of manpower.

So the medical board adjusted the clearinghouse mechanisms and gradually, doctors and hospitals came back to the central system. With tweaks over time, it is still in use, and over 99% of the hospitals and doctors use the current NIRMP matching system. The secret to their tweak? Well, you guessed it: to try to provide a stable match among the participants[2].

If this one evidence is not enough, let me show you an elegant table from Roth (2002) and pay particular attention to the United Kingdom's medical programs. UK has very localized medical intern programs, so different regions and cities used different matching systems. There were modifications and abandonments over time. Some programs survived and some programs did not. As

[2]Roth (1984) detailed the evolution and the tweaks.

TABLE I

STABLE AND UNSTABLE (CENTRALIZED) MECHANISMS

Market	Stable	Still in use (halted unraveling)
American medical markets		
NRMP	yes	yes (new design in '98)
Medical Specialties	yes	yes (about 30 markets)
British Regional Medical Markets		
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
Other healthcare markets		
Dental Residencies	yes	yes
Osteopaths (<'94)	no	no
Osteopaths (≥'94)	yes	yes
Pharmacists	yes	yes
Other markets and matching processes		
Canadian Lawyers	yes	yes (except in British Columbia since 1996)
Sororities	yes (at equilibrium)	yes

Table 1: Roth (2002)

you can see, all the mechanisms that yielded stable outcomes are still in use, but many of the unstable mechanisms went obsolete. A similar evolution is undergoing in the school choice programs. Let me just briefly say that the Boston School District has abandoned the "Boston Mechanism" and switched to the "Gale-Shapley algorithm" which we will discuss soon.

Let's recap what we have learned so far. We constructed a general and flexible model of the two-sided matching markets. We defined a stable and individually rational outcome in these matching markets: essentially no agent can ditch his or her current partner(s) and get a better match, and I tried to convince you that achieving stable outcome should be an essential goal in real world matching problems.

But does a stable outcome always exist? We answer this question next.

3 The Algorithm

Yes, a stable outcome always exists. We can prove the existence by constructing an algorithm that always returns a stable outcome. We now describe the algorithm. For simplicity, assume that all men are acceptable to a woman, and all women are acceptable to a man (w is acceptable to m if $wP_m\emptyset$), so any matching is individually rational and we can focus on achieving the stable outcome.

Theorem 1. *A stable match exists.*

The algorithm is as follows. First, every man proposes to his favorite woman. As a result, each woman receives many, one, or zero proposal. If a woman has more than one proposal, she rejects all but one man - the man she prefers the most among all the proposers. He is placed on a temporary acceptance list. Then each man who has been rejected and is not on a temporary acceptance list proposes to his second favorite woman. With the new proposers and possibly a man on the temporary acceptance list, each woman has many, one, or zero man to choose from. She again rejects all but one - the one she most prefers among the new proposers and the one on the temporary acceptance list. Each man who has been rejected proposes to his favorite woman who has not rejected him, and each woman rejects all but one. The process repeats. The process ends when no man proposes. The temporary acceptance list becomes permanent and this is the final matching which we denote by μ^*. This process terminates because there is a finite number of agents with each step at least one agent being rejected (suppose there are n men and n women, the process takes at most $n^2 - 2n + 2$ steps).

We can simply describe the process as follows.

Step 1: a) Each man proposes to his favorite woman,

b) Each woman keeps her favorite man among the men who propose to her and rejects all others.

Step k: a) Each man rejected in Step k-1 proposes to his most preferred woman among the ones who have not rejected him in the previous steps,

b) Each woman keeps her favorite man among the men who have proposed to her and rejects all others.

Terminates: No man proposes and all temporary acceptances become permanent.

Let's see how this algorithm finds our stable outcome in the prom example. Recall that Alex and Bob like Alice more, but Alice likes Alex more and Betty likes Bob more, and no one wants to go alone.

Step 1: a) Alex and Bob propose to Alice.

b) Alice has two proposals and Betty receives none: Alice rejects Bob and keeps Alex on the temporary acceptance list.

Step 2: a) Bob has been rejected in Step 1 and proposes to Betty, his second favorite, also the favorite among those who have not rejected him. Betty has only one offer and puts Bob on the temporary acceptance list.

Terminates: No man makes a proposal. The temporary acceptances become permanent: (Alex $\leftrightarrow$ Alice, Bob $\leftrightarrow$ Betty).

Now we need to verify that the outcome μ^* is indeed stable. Can there be a pair of m and w such that $w P_m \mu^*(m)$ and $m P_w \mu^*(w)$? The answer is no. Take any man m. All the women besides his

match $w = \mu^*(m)$ can be categorized into two types: 1) those who he prefers to w, $\{w' : w' P_m w\}$ and 2) those to whom he prefers w: $\{w' : w P_m w'\}$. We show no $w' \neq m$ can form a mutually beneficial pair with m. 1) If he prefers w' to w, it means that in the deferred acceptance algorithm, he has proposed to w' but is rejected. He is rejected because a better man m' has proposed to w', so $\mu^*(w') P_{w'} m$. Therefore, no woman from the first group wants to ditch her current mate and form a pair with m. 2) The second group is even easier: m does not want to ditch his current mate and form a pair with any woman in the second group, even if she wants. We see that for any $m \in M$, no $w \in W$ can form a profitable pair; therefore, no m and w can mutually improve upon μ^*, hence it is stable.

In the general many-to-many market, we modify the algorithm by allowing each man m to interact with at most q_m women (i.e. to propose to and to stay on the temporary acceptance list with at most q_m women) in a step and each woman to hold at most q_w men on her temporary acceptance list.

We apply it to the college admission example.

Step 1

a) α applies to B; β, γ, and δ apply to A.

b) A keeps β and γ, and rejects δ; B accepts α.

Step k

a) δ applies to B.

b) B rejects δ.

Terminates: No student proposes.

$$A \leftrightarrow (\beta, \gamma), B \leftrightarrow \alpha.$$

For the feature of the temporary acceptance list, this algorithm is named the deferred acceptance algorithm.

So far we have shown that a stable match exists, but in general, stable matches may not be unique. For example, suppose that Alex changed his preference: he likes Betty more than Alice now (all the others have the same preference),

$$\text{Betty } P_{\text{Alex}} \text{ Alice } P_{\text{Alex}} \emptyset$$

$$\text{Alice } P_{\text{Bob}} \text{ Betty } P_{\text{Bob}} \emptyset$$

$$\text{Alex } P_{\text{Alice}} \text{ Bob } P_{\text{Alice}} \emptyset$$

$$\text{Bob } P_{\text{Betty}} \text{ Alex } P_{\text{Betty}} \emptyset$$

Then both two-pairs matches are stable: (Alex $\leftrightarrow$ Alice, Bob $\leftrightarrow$ Betty), (Alex $\leftrightarrow$ Betty, Bob $\leftrightarrow$ Alice).

Let's look a little deeper into this stable match μ^* and the matching algorithm. Although it seems that women have some power in accepting and rejecting the candidate men, they are after all passively sought after. The men actually exhaustively search down their preference list. In fact, among all the stable matches, every man prefers the stable match achieved from the men-proposing deferred acceptance algorithm. For example, with this new preferences of the boys and girls, the stable match achieved by the deferred acceptance is (Alex $\leftrightarrow$ Betty, Bob $\leftrightarrow$ Alice) in which the boys each get their favorite girl.

Theorem 2. *Among all the stable matches, every man prefers the stable match obtained from the men-proposing deferred acceptance algorithm.*

We say that a woman w is **possible** for a man m if in some stable match μ, $w = \mu(m)$. In order to show the claim, we need to show that all the women who have ever rejected a man in the men-proposing deferred acceptance algorithm are impossible for a man. Since each man sequentially proposes from his preference list, his mate is the most preferred among all his possible women. The proof follows from induction. Suppose up to step $k-1$ of the deferred acceptance algorithm, no man has been rejected by a woman possible for him; in other words, the rejections, if any, were all made by women impossible for men. Now in step k, w rejects m' and keeps m. We need to show that w is impossible for m', so that in all steps of the algorithm, only women impossible for m' rejected him. Suppose otherwise: w is actually possible for m', so a hypothetical stable match μ' exists such that $m' = \mu'(w)$. Then the man m must be with his other possible woman $w' = \mu'(m)$ in μ'. But $wP_m w'$ because all the women better than w are not possible for m. On the other hand, $mP_w m'$, as the behavior of w rejecting m' for m indicates. Therefore, m and w can benefit from matching with each other by ditching their mates in μ', so μ' cannot be a stable match - a contraction so w is impossible for m'.

Furthermore, we can also show that the stable match from the men-proposing procedure is the least preferred by all women. If instead women propose, then the stable match is the most preferred stable match for all women. And if men-proposing and women-proposing outcomes coincide, then the stable outcome is unique.

Let's summarize what we have done so far. We demonstrated that a stable outcome always exists and we introduced an algorithm to systematically find a stable outcome. This matching outcome is special in the sense that men sequentially exhaust their options and get the best they could. Thus far, no game theory is involved. Even minimal math and math symbols are involved. We have kept the preferences to be known throughout the lecture, and we are now ready to relax this assumption and let agents report their preferences. A noncooperative game naturally arises.

4 Strategic Issues

Given the agents' preferences, the deferred acceptance algorithm spits out a stable match. However, in general, in practice, the market designer (the clearinghouse in the NRMP, for example) does not know the true preferences of the agents, and the agents must report these preferences. A mechanism therefore consists of two steps,

1. Let each agent report his or her preference
2. Based on the reported preferences, compute an outcome

In the deferred acceptance, the outcome computed is a stable matching. In the voting context, part 2 of the mechanism computes the winning candidate. The rest of the lecture concerns with the incentives of agents to report truthfully in the first part of the mechanism. Specifically, we want to answer the following question, if the men-proposing deferred acceptance algorithm is used to compute the stable match, does the agent always have the incentive to report his or her true preference? In other words, is the deferred acceptance algorithm strategyproof?

Agents reporting preferences becomes a game. In this game, each man m has strategy P_m and each woman w has strategy P_w. Let $\mu^*(m|P)$ denote m's mate from the DA algorithm when the preferences reported are $P = \left((P_m)_{m\in M}, (P_w)_{w\in W}\right)$. We say that the mechanism μ^* is strategyproof if it is a dominant strategy for all agents to report truthfully.

Definition 5. μ^* is **strategyproof** if for all $\widehat{P}$, for all $i \in M \cup W$,

$$\mu^*\left(i|P_i, \widehat{P}_{-i}\right) R_i \, \mu^*\left(i|\widehat{P}_i, \widehat{P}_{-i}\right).$$

Theorem 3. *The men-proposing deferred acceptance algorithm is not strategyproof.*

In order to show that the algorithm is not strategyproof, we only need to show under some preference reporting, an agent can profitably gain by misreporting. We return to our prom example, with the changed preferences, so that Alex's favorite is Betty, Betty's favorite is Bob, and Bob's favorite is Alice, and Alice's favorite is Alex. Suppose they all report truthfully. Alex proposes to Betty and Bob proposes to Alice. They both get accepted and the final match is (Alex-Betty, Bob-Alice). If someone can gain by reporting a preference other than the true preference P_i, then we show that the algorithm is not strategyproof. Indeed, there is a profitable manipulation. Consider Betty, instead of reporting true preference, Bob P_{Betty} Alex P_{Betty} $\emptyset$, she reports Bob P_{Betty} $\emptyset$ P_{Betty} Alex - she reports that Alex is unacceptable to her. Then when we run the algorithm, Alex is rejected by Betty in the first round. He then proposes to Alice who rejects Bob. Bob then proposes to Betty and gets accepted. The final match is then (Alex $\leftrightarrow$ Alice, Bob $\leftrightarrow$ Betty). Betty is strictly better off by partnering with Bob instead of Alex.

So how did Betty do that? She pretends to reject Alex initially, and Alex goes to Alice who dumps Bob, and Bob comes to Betty instead. This is like a high school drama, but theoretically, Betty creates a rejection chain that leads to a better match for her. Whenever such rejection chain exists, someone has incentive to misreport and reject one of his/her acceptable candidate to trigger the potential rejection chain.

You may ask, is there any strategyproof mechanism that returns a stable outcome? The answer is unfortunately no.

Theorem 4. *There does not exist a stable and strategyproof mechanism.*

The proof uses the same example. The two stable outcomes when everyone reports truthfully are (Alex $\leftrightarrow$ Alice, Bob $\leftrightarrow$ Betty) and (Alex $\leftrightarrow$ Betty, Bob $\leftrightarrow$ Alice). If the stable mechanism returns the (AA, BB) outcome, then Betty can misreport as above and gets a better outcome. If the stable mechanism returns the (AB, BA) outcome, then Alex can trigger a rejection chain by reporting Alice $P_{\text{Alex}}\,\emptyset\,P_{\text{Alex}}$ Betty, so that Alex gets Alice instead of Betty. Therefore, truth-telling cannot even be a Nash equilibrium so it cannot be a dominant strategy.

This impossibility result in some way resembles Gibbard-Satterthwaite impossibility theorem that there is no non-dictatorial and strategyproof mechanism. Maybe you are pessimistic, but don't be too so. We have a partially bright result. The agents who propose have no incentive to misreport their true preferences.

Theorem 5 (Dubins and Freedman, 1981)**.** *In the men-proposing deferred acceptance algorithm, it is a dominant strategy for all the men to reveal their true preferences.*

There is an active literature on the strategic issue when the market is large.

Theorem 6 (Lee, 2013)**.** *For every $\varepsilon > 0$, the expected proportion of men (and women) that have less than ε differences between utilities from the men-optimal match and the women-optimal match converges to one as the market size increases.*

5 Conclusion

In this lecture, we studied the two-sided matching problems. We introduced the concept of stability and studied the original form of the Gale-Shapley algorithm that guarantees to find a stable outcome. The agents' preferences over the agents on the other side of the market are given and public information. We relaxed this environment by studying the noncooperative game of reporting these preferences when the algorithm needs private inputs. The results are not ideal but not entirely pessimistic: the agents who propose will always report truthfully, but the agents who receive proposals may have profitable manipulations.

A huge literature has sprouted in the past fifty years, and the literature has particularly focused on the strategic interactions this algorithm and extensions and variants of the algorithm may induce.

References

Dubins, Lester E. and David A. Freedman, "Machiavelli and the Gale-Shapley Algorithm," *The American Mathematical Monthly*, August - September 1981, *88* (7), 485–494.

Lee, SangMok, "Incentive Compatibility of Large Centralized Matching Markets," January 2013. Mimeo.

Roth, Alvin E., "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of Political Economy*, 1984, *92* (6), pp. 991–1016.

__ , "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics," *Econometrica*, 2002, *70* (4), 1341–1378.

Exercises

Exercise 1. Find all the stable matchings of the Prom Example 1, the College Admission Example, and the Prom Example 2.

Exercise 2. Find the men-optimal and women-optimal stable matchings under the following rank matrix. $\alpha, \beta, \gamma, \delta$ are men and A, B, C, D are women, and the rank matrix is interpreted as follows: α ranks A first, B second, C third, and D fourth.

	A	B	C	D
α	$1,3$	$2,2$	$3,1$	$4,3$
β	$1,4$	$2,3$	$3,2$	$4,4$
γ	$3,1$	$1,4$	$2,3$	$4,2$
δ	$2,2$	$3,1$	$1,4$	$4,1$

Exercise 3. Suppose there are n men and n women in the one-to-one matching market. Show it takes at most $n^2 - 2n + 2$ steps for the deferred acceptance algorithm to terminate. Explicitly show preferences of 3 men and 3 women such that the algorithm takes $3^2 - 2 \times 3 + 2 = 5$ steps under these preferences.

Exercise 4. Under the three examples in the lecture (the Prom Example 1, the College Admission Example, and the Prom Example 2), will agents have incentives to misreport if the men-proposing algorithm is run? In other words, is it a Nash equilibrium for everyone to report truthfully if everyone knows the others' preferences? Prove so or show otherwise by identifying a profitable manipulation.