Treball final de grau

GRAU DE MATEMÀTIQUES

Facultat de Matemàtiques i Informàtica
Universitat de Barcelona

Two-Sided Matching Markets

Autora: Paula Fernández Grande

Director: Dr. Javier Martínez de Albéniz
Realitzat a: Departament de Matemàtica Econòmica, Financera i Actuarial

Barcelona, 23 de juny de 2018

Abstract

Two-sided matching markets are the main object of study of Matching theory, a branch of Game theory that involves discrete mathematics. A two-sided matching market is a market whose agents belong to one of two distinct sets and such that its outcome is an assignment of agents from one set to the other, and vice-versa.

In this work, we first give some basic definitions of Game theory, and turn then to study two kinds of matching markets: one-to-one matching markets, in which an agent from one set ends up matched to at most one agent on the other set, and many-to-one matching markets, in which agents on one of the two sets might end up matched to several agents on the other set. The former will be studied by means of the marriage problem in Chapter 3, and the latter by means of the college admissions problem in Chapter 4. To conclude, in Chapter 5 there is an outline of some alternative mechanisms used in different contexts to solve problems that are similar to those already studied.

2010 Mathematics Subject Classification. 91A40, 91B68.

Acknowledgements

I would first like to thank my director Javier Martínez de Albéniz for all the advice and help throughout these months. Secondly, I would like to thank Maite, Paco and Martí for all their support, I would not have got here without it. And last but not least, I would like to thank Pep Lamarca: gràcies per ensenyar-me a domesticar dracs.

Contents

Chapter 1

Introduction

The research that led to the analysis of two-sided matching markets has its roots in the analysis of markets using the tools of linear programming and game theory. Markets with indivisible goods are analysed firstly in the classical models of Gale (1960) [12] with the model of n houses bought by n buyers, in Chapter V; later Gale and Shapley (1962) [14] with the college admissions model and finally Shapley and Shubik (1972) [33] with the study of the assignment game with quasilinear utility.

Different variants of these models have many applications in the theory of recruitment systems, auction markets, labour markets and so on (see, e.g., Roth and Sotomayor (1990) [29]; Crawford and Knoer (1981) [7]). Many interesting applications are presented in the survey of Sönmez and Ünver (2014) [34] and the role of such models in modern economics is explained in Roth (2002) [30].

The first appearance of two-sided matching markets was in 1962 with the paper of David Gale and Lloyd Shapley [14], where they presented matching markets as mathematical problems that could be solved using an algorithm they had designed.

The aim of this theory is to understand and set formal and mathematical foundations to two-sided matching markets. In these markets, we assume that there are two sides (firms and workers, jobs and workers, organ donors and recipients, etc.) and an exchange is made without prices; let us start by defining more clearly what this all means. First of all, the term *two-sided* refers to the fact that agents on these markets are clearly divided into two disjoint sets, unlike agents on commodity markets, where the market price may determine whether someone buys or sells, and so one can decide to change sides at any time. Secondly, the term *matching* refers to the bilateral exchange that is produced in these markets; this is also different from commodity markets, in which chains of exchanges can be made and the good you end up buying may not be the one brought to the market by the agent who sold it to you.

A clear example of two-sided matching markets are labour markets, in which workers seek a job and employers need workers. In this case, one side of the market is made up of workers and the other one of companies; and obviously, if a company employs John, then John works for that company, but John cannot become a company and the company will never become an employee.

Another important characteristic of these markets is that money is not always the means of exchange. In fact, it is often considered unethical to use money in these markets, or even illegal. For instance, it is clear that selling or buying organs is not the ideal Sunday morning plan.

But, is this kind of market that common? Are they so different from commodity markets that they cannot be interpreted in the same way? Well, the answers are fairly straightforward: two-sided matching markets are very common and they are indeed different from commodity markets.

According to Gale (2001) [13], it all started with an article on The New Yorker magazine on 10 September 1960, whose author had been investigating how the admissions of undergraduate students office worked in Yale University, and what he found was absolute chaos. That was when Gale, together with Shapley, started to study the problem and wrote the paper we have mentioned above, creating a field of study which has not ceased to expand until today. In this paper, they introduced the marriage problem, which we will explain in Chapter 3, and the college admissions problem, which we will explain in Chapter 4; they also described an algorithm to find a solution to these problems, the Deferred Acceptance Algorithm (DAA). Years later, it was discovered that equivalent algorithms had already been built independently several times. Thus, these problems had already arisen before and had been tackled, sometimes successfully, despite the lack of theoretical work supporting them.

Actually, it was not until 1984 that Alvin E. Roth found out that an equivalent algorithm to the DAA had been in use since the 1950's by the NIMP (National Intern Matching Program) to assign senior medical students to the hospitals where they would undertake their internships.[1] The original algorithm proposed in the beginning of the century did not have the desired properties, and the matching resulting from it had a number of problems, resulting in the market not working well; in 1945, the market eventually collapsed (many students had their internship sealed almost two whole years before it began). They tried to solve this problem in a series of ways, until 1952, when the algorithm implemented was successful. In fact, even though students were permitted to find internships outside the system, that year over 95% of them preferred to use the NIMP algorithm. This was so until the 70's, when women started to become a greater percentage of the students of universities, and many couples made up of two students looking for an internship together saw that they could get a better position by finding their internships by themselves.[2] In this case, it has been shown that there might not exist any solution with the most desired property, which is stability.[3]

However, this is only one of the many real-life examples of two-sided matching markets we could outline: there is a vast variety of examples, from college admissions to kidney exchange or dating websites. That is why, in 2012, Alvin E. Roth and Lloyd S. Shapley were awarded the Nobel Prize in Economic Sciences, "for the theory of stable allocations and the practice of market design".[4]

To conclude, this work is mainly focused on the study two of the most famous two-sided matching markets, conceiving them as problems to solve: the marriage problem

[1] This case and its history are extensively explained in [24].
[2] We will not tackle this problem in our work, but it is studied in [19].
[3] We will introduce this concept in Chapter 3.
[4] See [21].

and the college admissions problem. In Chapter 2, we will outline some basic concepts of Game theory that will be used in the following chapters. In Chapter 3, we will study the marriage problem, explaining the formal model and some results, and introducing the DAA as an algorithm to find a possible solution for this model. In Chapter 4, we will study the college admissions problem; we will first outline the model, and as it is (kind of) a generalisation of the marriage problem, some of the results that are not proved in Chapter 3 are proved in Chapter 4, while some more observations and results are added. In Chapter 5 we will give a quick rundown on some of the different algorithms that are worldwide used to find solutions to this kind of problems, each with a brief explanation on its uses and hypothesis.

Chapter 2

Game theory: the essentials

In this brief chapter, we introduce the central concepts of game theory. These concepts will be needed to understand the strategic aspects of matching markets, and are essential to grasp the meaning of situations where agents act without supervision in a known setting. The main issue is the concept of Nash equilibrium.

Plainly speaking, a *game* is a situation governed by rules and with a well-defined outcome, characterised by strategic interdependence (Gardner, 1996) [16].

Definition 2.1. The *normal form* of a *game* G is formed by three elements:

(i) a finite set of *players* $\mathcal{I} = \{1, \ldots, n\}$,

(ii) $\forall i \in \mathcal{I}$, a set of *strategies* S_i, and

(iii) $\forall i \in \mathcal{I}$, a *payoff function* $u_i : S \to \mathbb{R}$, where $S = S_1 \times \ldots \times S_n$, also called *utility function*.

Thus, G is defined by the triple $(\mathcal{I}, S, u)$, and the ultimate objective of every $i \in \mathcal{I}$ is to maximize her utility u_i, which depends on the election made not only by herself but by all players.

We call every element $s = (s_1, \ldots, s_n) \in S$ a *strategy profile*. Payoff functions are a way of ordering players' preferences over different outcomes of the game. Clearly, if $u_i(s) < u_i(s')$, player i will prefer the strategy profile s' over s. Here is where strategic interdependence is fundamental.

Given a player $i \in \mathcal{I}$ and a strategy profile $s \in S$, we denote by s_{-i} the strategy profile of all players except player i, that is, $s_{-i} \in \prod_{j \in \mathcal{I} \setminus \{i\}} S_j =: S_{-i}$. Therefore, $s = (s_{-i}, s_i)$, and we are able to assert that, when playing the game $G = (\mathcal{I}, S, u)$, the objective of player $i \in \mathcal{I}$ is to maximize her gains given the strategy profile of all other players, that is to find, given s^*_{-i},

$$\arg\max_{s_i \in S_i} u_i(s^*_{-i}, s_i).$$

This is generally a set, called *best response of i to s^*_{-i}*, and it is denoted by $BR_i(s^*_{-i})$.

Definition 2.2. Let G be the game $(\mathcal{I}, S, u)$. Given the strategies $s_i, s_i' \in S_i$ of a player $i \in \mathcal{I}$, strategy s_i' is *strictly dominated by* strategy s_i if $u_i(s_{-i}, s_i) > u_i(s_{-i}, s_i')$, $\forall s_{-i} \in S_{-i}$. In this case, s_i is a *dominant strategy*. If no such s_i exists, the strategy s_i' is called an *undominated strategy*.

Definition 2.3. Let $G = (\mathcal{I}, S, u)$ be a game. Given two strategies of player $i \in \mathcal{I}$, $s_i, s_i' \in S_i$, strategy s_i' is *weakly dominated by* strategy s_i if $u_i(s_{-i}, s_i) \geq u_i(s_{-i}, s_i')$, $\forall s_{-i} \in S_{-i}$.

The structure of a game G is often considered to be of *common knowledge*. This means that every $i \in \mathcal{I}$ knows S and u, and that every other player knows that she knows that, and so on ad infinitum. That is, not only i knows S_j and u_j, $\forall j \in \mathcal{I}$, but she also knows that j knows S_i and u_i, and vice-versa. If this is so, G has *complete information*.

Plus, we assume players to be rational, meaning they will not choose dominated strategies, and that this is also common knowledge, so they know others won't either (and know everybody knows this, etc.).

The following concept regarding strategies will be used in further chapters too.

Definition 2.4. Let G be the game $(\mathcal{I}, S, u)$ and let $s, s' \in S$ be two strategy profiles. Strategy s *Pareto-dominates* s' if

$$\forall i \in \mathcal{I}, u_i(s) \geq u_i(s'), \text{ and } \exists j \in \mathcal{I} \text{ such that } u_j(s) > u_j(s').$$

Moreover, s is *Pareto-optimal* (or *Pareto-efficient*) if it is not dominated by any other strategy $s' \in S \backslash \{s\}$.

Observe that whilst definitions 2.2 and 2.3 refer to the strategies of a single player, definition 2.4 refers to strategy profiles as a whole, and thus they define completely different concepts.

Let us see an example of dominance and Pareto-dominance. This example is quite common in our daily lives, although probably not in the context it is here presented.

Example 2.5. (The prisoners' dilemma) There are two suspects of having committed both a murder and a robbery, who are put into separate cells. The police have solid proof of the robbery, but not of the murder. They both have a chance to confess to murdering. If one does and the other does not, the one who has confessed incriminates the other, and the result is that the one who has confessed does not go to jail, whilst the one who has not goes to jail for 15 years; if the two of them confess, both go to jail for 10 years; finally, if neither of them confesses they only go to jail for 1 year.

We have the game $G = (\mathcal{I}, S, u)$, with: $\mathcal{I} = \{1, 2\}$, $S_1 = S_2 = \{C, NC\}$, where C stands for "confess" and NC for "not confess", and $S = S_1 \times S_2$, $u_1 : S \to \mathbb{R}$, such that $u_1(C, C) = -10$, $u_1(C, NC) = 0$, $u_1(NC, C) = -15$, $u_1(NC, NC) = -1$, and $u_2 : S \to \mathbb{R}$, such that $u_2(C, C) = -10$, $u_2(C, NC) = -15$, $u_2(NC, C) = 0$, $u_2(NC, NC) = -1$.

The following table is the best way to see all these payoffs clearly; for each strategy profile, it gives us the utility of player 1 in the first place, and then the utility of player 2:

1 \ 2	C	NC
C	$-10; -10$	$0; -15$
NC	$-15; 0$	$-1; -1$

Considering only the strategies of a single player, it is clear that the best she can do, regardless of whether the other suspect has confessed or not, is to choose strategy C. That means that NC is dominated by C, as any rational player would never choose NC over C. However, the payoff for both players if they choose their dominant strategy C is -10, which is obviously worse than the payoff they would get if they both choose NC, -1. Thus, the strategy profile (NC, NC) Pareto-dominates the strategy profile (C, C).

Therefore, we have seen that Pareto-dominance and dominance might be conflicting concepts. This implies that neither of them can be an acceptable solution to the game. The following definition is deemed the most important of the many solution concepts that there exist, and it is based on the fact that players of the game $(\mathcal{I}, S, u)$ behave rationally; this concept of solution yields a strategy profile $s^* = (s_1^*, \ldots, s_n^*) \in S$ such that no player $i \in \mathcal{I}$ is able to improve her situation by choosing another strategy $s_i \in S_i$, with $s_i \neq s_i^*$. In this case, s_i^* is called i's *best response*, as it is the best strategy she can choose given the strategies of everyone else, s_{-i}^*.

Definition 2.6. Let $G = (\mathcal{I}, S, u)$ be a game. A *Nash equilibrium* of G is a strategy profile $s^* = (s_1^*, \ldots, s_n^*) \in S$ such that, $\forall i \in \mathcal{I}$ and $\forall s_i \in S_i$,

$$u_i(s^*) \geq u_i(s_{-i}^*, s_i)$$

This concept was first introduced by John Nash (1950) [20], and it is a key concept of game theory. Notice that the existence of a Nash equilibrium is not guaranteed. Nevertheless, Nash (1950) [20] showed that under certain conditions, at least one Nash equilibrium does exist. It is not a constructive result.

Observe that in Example 2.5, the strategy profile (C, C) is a Nash equilibrium.

Chapter 3

The marriage problem

In this chapter, we study the marriage problem. This problem falls into the category of two-sided matching markets. As we have said before, a two-sided matching market is one in which agents are clearly divided into two disjoint sets, in such a way that no agent can decide to change sides and become part of the other set; the final outcome of the market is an assignment between the two sets. In particular, the marriage problem is a case of one-to-one matching, because every agent ends up matched to at most one agent of the other set.

The reason why it is called the marriage problem is because of the natural parallelism: there is a group of women and a group of men, and the aim is for them to marry (heterosexually and monogamously). This problem was first introduced by Gale and Shapley (1962) [14], and has since been widely studied. It is the first step towards understanding more complex cases, such as the college admissions problem, in which agents on one side of the market might be assigned many agents on the other side. We will study such markets in the next chapter.

3.1 The model

Formally, there are two finite and disjoint sets: $W = \{w_1, \dots, w_n\}$ is the set of women and $M = \{m_1, \dots, m_p\}$ is the set of men. Each woman w has preferences over $M \cup \{w\}$ and each man m has preferences over $W \cup \{m\}$; this means that any agent may prefer to stay single than to be matched to some (or even any) of the agents on the other side. These preferences are represented by ordered lists; for example, if the preferences of woman w are $P(w) = m_2, m_7$, then w prefers m_2 to any other man, her second choice is m_7, and she prefers to remain single rather than to be matched to any of the other available men. Implicitly, we have defined for every woman $w \in W$, a total order $>_w$ in the set $M \cup \{w\}$, so that $m >_w m'$ means w prefers m to m'. Analogously, $>_m$ is a total order in the set $W \cup \{m\}, \forall m \in M$. The set of all these total orders is $> := (>_x)_{x \in M \cup W}$. We will call $>_x$ indistinctly agent x's *preference* relation or *order* relation.

We will stick to the case in which preferences are strict, so no individual is indifferent when faced with two choices.[1] We do so because our main purpose is the college admissions problem, in which both students and colleges are to submit a strict list of preferences; another reason to do this is that if we relax this assumption and consider non-strict order relations, some of the results do not hold. Moreover, from now on, we will consider all agents to be rational, meaning their preferences are complete (any two options can be compared) and transitive. We will also say that man m is *acceptable* to woman w if $m >_w w$; analogously, w is *acceptable* to man m if $w >_m m$.

We denote a preference profile by $P = P(m_1) \times \ldots \times P(m_p) \times P(w_1) \times \ldots \times P(w_n)$, where $P(x)$ is the preference list submitted by $x \in M \cup W$. For now, we will assume that every $x \in M \cup W$ submits a preference list with accordance to her preference relation $>_x$; however, we will see later that agents may have incentives not to do so. The set of all preference profiles is $\mathcal{P}(M, W)$.

We now have all the ingredients of a marriage problem: a specific marriage problem or marriage market is denoted by the triple $(M, W; >)$. The problem we are faced with is the following: given the preferences submitted by every individual, what kind of outcome will arise from their collective behaviour?

Definition 3.1. Given the sets of men M and women W, a *matching* is a one-to-one correspondence $\mu : M \cup W \longrightarrow M \cup W$ such that

(a) $\mu^2(x) = x$, $\forall x \in M \cup W$,

(b) $\mu(w) \in M \cup \{w\}$, $\forall w \in W$, and

(c) $\mu(m) \in W \cup \{m\}$, $\forall m \in M$.

The set of all matchings is denoted by $\mathcal{M}(M, W)$.

Example 3.2. Let $M = \{m_1, m_2, m_3\}$ and $W = \{w_1, w_2, w_3\}$ be sets of men and women respectively. Then, the matching represented by

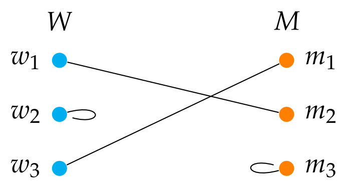
$$\mu = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_2 & (w_2) & m_1 \end{pmatrix}$$

matches w_1 to m_2 and w_3 to m_1, whereas w_2 remains *self-matched* (that is, single).

Notice that because of the fact that a matching μ must be of order two, if woman w is matched to man m, then m is matched to w.

Matchings can also be represented using graphs. These graphs will have two distinct sets of nodes, but need not be bipartite graphs, because there can be loops.

Example 3.3. The matching of Example 3.2 can be represented using the following graph.



[1]That is, $\forall w \in W$ and $x_1, x_2 \in M \cup \{w\}$, either $x_1 >_w x_2$ or $x_2 >_w x_1$; similarly, we are assuming that $\forall m \in M$ and $x_1, x_2 \in W \cup \{m\}$, either $x_1 >_m x_2$ or $x_2 >_m x_1$.

The representation through graphs is more visual and all the information is explicit, while the usual representation gives implicit information; for instance, in Example 3.2, it is clear that w_2 is self-matched, and so is man m_3 even though he does not appear.

From our definition, we can define a preference relation over $\mathcal{M}(M, W)$ for each agent. We assume that the preferences of individuals over matchings concern only their own match, and thus we will write $\forall x \in M \cup W, \forall \mu, \nu \in \mathcal{M}(M, W), \mu \succ_x \nu \iff \mu(x) >_x \nu(x)$. What is more, x is indifferent between two different matchings μ and ν only if she has the same partner at both, that is, $\mu \succeq_x \nu \iff \mu(x) >_x \nu(x)$ or $\mu(x) = \nu(x)$.[2]

We write $\mu \succeq_W \nu$ to denote that every woman in W likes μ at least as well as she likes ν, and $\mu \succ_W \nu$ to denote that $\forall w \in W$, $\mu \succeq_w \nu$ and $\exists w \in W$ such that $\mu \succ_w \nu$. Clearly, $\succeq_M$ and $\succ_M$ are defined analogously.

3.2 Stability

Since our current purpose is to observe which outcomes of the market are more likely to occur, we must first state the rules of this game, in order to rule out some inconvenient matchings. The general rules of the marriage problem are simple: any couple that agrees to marry may do so, and every individual is free to remain single if she wishes.

Definition 3.4. A matching $\mu \in \mathcal{M}(M, W)$ is *individually irrational*, or it is *blocked by an individual*, if $\exists x \in M \cup W$ such that $x >_x \mu(x)$.

Example 3.5. Consider the marriage market $(M, W; >)$, where $M = \{m_1, m_2, m_3\}$ and $W = \{w_1, w_2, w_3\}$ are the sets of men and women, respectively, endowed with the following preferences:

$$\begin{aligned} P(w_1) &= m_1, m_2; & P(m_1) &= w_1, w_2; \\ P(w_2) &= m_2, m_1, m_3; & P(m_2) &= w_2, w_1, w_3; \\ P(w_3) &= m_3, m_1; & P(m_3) &= w_3, w_2, w_1. \end{aligned}$$

In this case, the matching

$$\mu = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_2 & m_3 & m_1 \end{pmatrix}$$

is individually irrational, since $\mu(m_1) = w_3$, and $m_1 >_{m_1} w_3$, i.e. w_3 is *unacceptable* to m_1.

Thus, an individually irrational matching is not a feasible outcome of the market, since any individual not matched to an acceptable mate would have done better by remaining single, an option that is always available. Therefore, all feasible outcomes must be *individually rational*.

Observe that for all $P \in \mathcal{P}(M, W)$, there is an individually rational matching, the one in which everybody is self-matched.

Now, let us consider blocking couples as well:

Definition 3.6. A couple $\{m, w\} \subset M \cup W$ *blocks* a matching $\mu \in \mathcal{M}(M, W)$ if $m >_w \mu(w)$ and $w >_m \mu(m)$.

[2] As different matchings may result in an individual matched to the same partner, in these preference relations there is indeed the possibility to be indifferent between two outcomes.

That is, we have a matching μ so that w and m both prefer to be matched to each other than to their current partners. In this case, they would not be breaking any rule if they both refused their match and married each other.

Example 3.7. Recalling the marriage market of Example 3.5, the pair $\{m_2, w_2\}$ blocks the matching μ considered.

If preferences are common knowledge, this kind of matchings is not likely to occur either, and the only remaining feasible outcomes are the ones that meet the following definition.

Definition 3.8. Let $(M, W; >)$ be a marriage market. Then $\mu \in \mathcal{M}(M, W)$ is a *stable matching* if it is individually rational and it is not blocked by any pair of agents.

The set of all stable matchings of the marriage problem $(M, W; >)$ is denoted by $\mathcal{S}(P)$.

Notice that when preferences are not common knowledge, there might exist an *unstable matching*[3] that occurs despite there being a pair $\{m, w\} \subset M \cup W$ that blocks it, only because m and w were unaware of each other's preferences. However, we will henceforth consider markets in which these problems do not emerge, and thus if possible, it is desirable for the outcome to be in $\mathcal{S}(P)$.

We will now show that for any $(M, W; >)$, there always exists a stable matching. We will do so through a *mechanism*, which is an algorithm that builds an output -a matching- from the input $(M, W; >)$.

Theorem 3.9. (Gale and Shapley, 1962) *For any marriage problem $(M, W; >)$, $\mathcal{S}(P) \neq \emptyset$.*

Proof. We will display an algorithm that builds an outcome in $\mathcal{S}(P)$, starting from an arbitrary preference profile $P \in \mathcal{P}(M, W)$.

Step 1. (a) Each woman proposes to her most preferred man.

(b) Each man rejects all but his most preferred of the women who have proposed to him, and holds the chosen one as his suitor. If there is a man who receives no proposals, he remains provisionally self-matched.

⋮

Step k. (a) Each woman who has been rejected in the step $k-1$ proposes to the most preferred of those men who have not yet rejected her. If no such man exists, she remains self-matched.

(b) Each man holds as his suitor the woman he prefers among those who have proposed to him in his step, together with the one he had held as his suitor at the end of step $k-1$, and rejects the rest.

The algorithm stops after the first step in which no woman is rejected, or when every woman rejected has no more acceptable men. It is clear that this will happen eventually, since $|M \cup W| < \infty$; and obviously, the outcome is individually rational, because nobody

[3]Obviously, an unstable matching is a matching $\mu \in \mathcal{M}(M, W) \backslash \mathcal{S}(P)$.

is ever even temporarily matched to an unacceptable mate. The ones who remain single are the women rejected by all of their acceptable men and the men who have not received any proposal from an acceptable woman.

To see that the resulting matching is in $\mathcal{S}(P)$, let us suppose it is not. Then, as we have already ruled out the possibility of individual irrationality, we can only assume that there is a pair $\{m, w\}$ that blocks the matching. In this case, man m must be acceptable to woman w, and thus she must have proposed to him before proposing to her current match, and if this is so, he must have rejected her in favour of a preferred mate. Now, following from the transitivity of $>_m$, m prefers his match to w, and $\{m, w\}$ cannot be a blocking pair. □

This mechanism, called the Deferred Acceptance Algorithm (DAA), was proposed by Gale and Shapley (1962) [14]. The idea behind this name is that men can keep the best available woman at any step engaged, without taking any final decision yet. Notice that there is a symmetric algorithm in which men take the proposing role.

Example 3.10. Consider the marriage problem $(M, W; >)$, where $M = \{m_1, m_2, m_3\}$ is the set of men and $W = \{w_1, w_2, w_3, w_4\}$ is the set of women. Let the preference profile P be given by

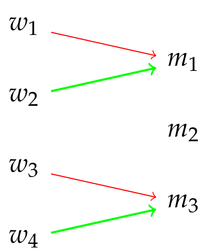
$$\begin{array}{ll} P(w_1) = m_1, m_3, m_2; & P(m_1) = w_2, w_3, w_1, w_4; \\ P(w_2) = m_1, m_2, m_3; & P(m_2) = w_3, w_4, w_1, w_2; \\ P(w_3) = m_3, m_2, m_1; & P(m_3) = w_2, w_1, w_4, w_3; \\ P(w_4) = m_3, m_2, m_1. & \end{array}$$

Let us apply the DAA.

Step 1. Both w_1 and w_2 propose to m_1, while w_3 and w_4 propose to m_3.

- $w_2 >_{m_1} w_1$, so m_1 rejects w_1 and keeps w_2 as his suitor.
- m_2 receives no proposals, and therefore remains provisionally self-matched.
- $w_4 >_{m_3} w_3$, so m_3 rejects w_3 and keeps w_4 as his suitor.

The following diagram shows every proposal that has been made in the first step; those that have been rejected are represented by a red edge and those that have been held are represented by a green one.

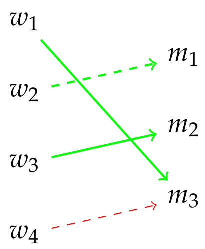


Step 2. Only w_1 and w_3, who have been rejected in the first step, make proposals. w_1 proposes to m_3 and w_3 proposes to m_2.

- m_1 does not receive any new proposals, so he holds w_2 as his suitor.

- m_2 holds w_3 as his suitor.
- $w_1 >_{m_3} w_4$, so m_3 rejects w_4 and keeps w_1 as his suitor.

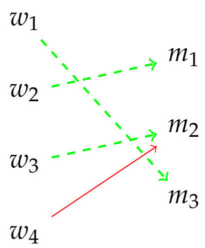
The following diagram shows the proceeding of step 2: dashed edges represent the tentative matches from step 1, and as in the diagram above, green edges represent held proposals and red ones, rejected proposals.



Step 3. Only w_4 has been rejected in step 2, so she is the only one proposing, and she proposes to m_2.

- m_1 keeps w_2 as his suitor.
- $w_3 >_{m_2} w_4$, so m_2 rejects w_4 and keeps w_3 as his suitor.
- m_3 keeps w_1 as his suitor.

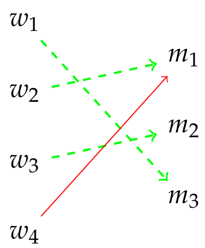
The diagram of step 3 is the following.



Step 4. Only w_4 has been rejected in step 3, so she is the only one proposing, and she proposes to m_1.

- $w_2 >_{m_1} w_4$, so m_1 rejects w_4 and keeps w_2 as his suitor.
- m_2 keeps w_3 as his suitor.
- m_3 keeps w_1 as his suitor.

The diagram of step 4 is the following.



w_4 has no more possible proposals to make, and so her only option is to remain single. Therefore, the algorithm ends after step 4, and every tentative match becomes definitive. The resulting matching is

$$\mu = \begin{pmatrix} w_1 & w_2 & w_3 & w_4 \\ m_3 & m_1 & m_2 & (w_4) \end{pmatrix}.$$

In our example, women take the proposing role. We will call the resulting matching from the women-proposing procedure μ_W, while the one resulting from the men-proposing procedure will be called μ_M. Both μ_W and μ_M are stable matchings, and will not usually be the same. There are many interesting results regarding μ_M and μ_W. For instance, the fact that $\mu_M \succeq_M \mu_W$ and $\mu_W \succeq_W \mu_M$.

Definition 3.11. Given a marriage market $(M, W; >)$, a stable matching $\mu \in \mathcal{S}(P)$ is *W-optimal* if $\mu \succeq_W \nu$, $\forall \nu \in \mathcal{S}(P)$.

An *M-optimal* matching is defined analogously. Gale and Shapley proved that such stable matchings actually exist, and what is more, they are unique. In fact, the W-optimal matching is μ_W and the M-optimal stable matching is μ_M.

Example 3.12. Let us consider the marriage market $(M, W; >)$, with $M = \{m_1, m_2, m_3\}$, $W = \{w_1, w_2, w_3\}$, and submitted preferences given by

$$\begin{array}{ll} P(w_1) = m_2, m_3, m_1; & P(m_1) = w_3, w_2, w_1; \\ P(w_2) = m_1, m_2, m_3; & P(m_2) = w_1, w_3, w_2; \\ P(w_3) = m_3, m_1, m_2; & P(m_3) = w_2, w_1, w_3. \end{array}$$

There are six possible matchings in which nobody remains single (and clearly, all of them are individually rational because nobody is unacceptable for anybody of the opposite sex):

$$\mu_1 = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_1 & m_2 & m_3 \end{pmatrix}, \quad \mu_2 = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_1 & m_3 & m_2 \end{pmatrix}, \quad \mu_3 = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_2 & m_1 & m_3 \end{pmatrix},$$
$$\mu_4 = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_2 & m_3 & m_1 \end{pmatrix}, \quad \mu_5 = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_3 & m_1 & m_2 \end{pmatrix}, \quad \mu_6 = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_3 & m_2 & m_1 \end{pmatrix}.$$

It is easy to check that μ_1 and μ_2 are blocked by the pair $\{m_2, w_1\}$, μ_5 is blocked by the pair $\{m_1, w_3\}$, and μ_6 is blocked by the pair $\{m_3, w_3\}$; μ_3 and μ_4 have no blocking pairs. Thus, $\mathcal{S}(P) = \{\mu_3, \mu_4\}$. Trivially, $\mu_W = \mu_3$ because every woman is matched to her most preferred mate, and similarly $\mu_M = \mu_4$.

So far, we have seen that the set of stable matchings is not empty; we have also begun to see that agents on opposite sides of the market have conflicting interests over the different stable outcomes, whilst agents on the same side may have common interests regarding the best stable matching they can achieve.

We have already said that the matching μ_W is the best stable matching the women can achieve, but is there any unstable matching in which they would all do better? In other words, are they paying a price for stability? The answer is negative.

Theorem 3.13. (Weak Pareto optimality for the women). *There is no individually rational matching $\mu \in \mathcal{M}(M, W)$ such that $\mu \succ_w \mu_W$, $\forall w \in W$.*

The proof of this theorem is by means of the deferred acceptance algorithm.[4] Obviously, the analogous result for the men is also true. Let us see a seemingly trivial example.

Example 3.14. Let $(M, W; >)$ be a marriage market with $M = \{m_1, m_2, m_3\}$, $W = \{w_1, w_2, w_3\}$ and the following submitted preferences:

$$\begin{array}{ll} P(m_1) = w_1, w_2; & P(w_1) = m_2, m_3; \\ P(m_2) = w_2, w_3; & P(w_2) = m_3, m_2; \\ P(m_3) = w_3, w_1; & P(w_3) = m_2, m_1. \end{array}$$

There is no individually rational matching in which nobody remains self-matched. If we apply the DAA with women proposing, we obtain the matching

$$\mu_W = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_3 & m_2 & (w_3) \end{pmatrix},$$

and no other matching gives every woman a strictly preferred matching. However, the following matching is weakly preferred by women to μ_W:

$$\mu = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_2 & m_3 & (w_3) \end{pmatrix}.$$

Indeed, $\mu \succ_{w_1} \mu_W$, $\mu \succ_{w_2} \mu_W$, and $\mu(w_3) = \mu_W(w_3)$.

As we have observed in this example, stability and weakly Pareto-optimality need not happen simultaneously. This is a crucial issue when trying to solve these problems in practice, as we have to decide which one of these two properties is more desirable. In Chapter 5 we will see how different mechanisms yield matchings with these different properties.

Prior to the following section, let us recall the previous chapter. We can now consider the marriage problem $(M, W; >)$ as a game $(\mathcal{I}, S, u)$ with the set of players being $\mathcal{I} = M \cup W$, the set of strategies $S = \mathcal{P}(M, W)$, and $u = (u_x)_{x \in M \cup W}$, with $u_x : S \to \mathbb{R}$ any function representing the utility of an agent $x \in M \cup W$, that is, such that $\forall s, s' \in S$, $u_x(s) > u_x(s') \iff x$ prefers the outcome when the strategy profile is s than the outcome when the strategy profile is s'.

3.3 Dominated matchings and core of the game

In this section, we will study the structure of the set of stable matchings $\mathcal{S}(P)$. But first, we need some previous definitions of general matching theory, such as dominated matchings or the core of a game.

So far, we have defined a marriage market $(M, W; >)$ as a game with the set of players $M \cup W$; we have also specified the possible outcomes, $\mathcal{M}(M, W)$, and its rules. These rules, together with the preference profile of the players $P \in \mathcal{P}(M, W)$, induce yet another relation on the set $\mathcal{M}(M, W)$, called dominance.

[4] See [29], Chapter 2.

Definition 3.15. Let $(M, W; >)$ be a marriage market and $\mu, \nu \in \mathcal{M}(M, W)$ two matchings. We say that ν is *dominated by* μ if there is a subset $S \subset M \cup W$, namely a coalition of agents, such that

(a) $\forall x \in S$, $\mu \succ_x \nu$, and

(b) the rules of the market give S the power to impose μ over ν.[5]

In view of this, if ν is dominated by μ, we can be quite sure that ν will not be the outcome of the game, because there is a set of players who have the power to impose μ, an outcome they all prefer to ν. Thus, dominance seems to be somewhat related to stability.

However, notice that until now, we had only considered dominance via single individuals, ruling out individually irrational outcomes, and via couples $\{m, w\} \subset M \cup W$, eliminating every outcome blocked by pairs $\{m, w\}$ with $m \in M$ and $w \in W$. However, we had not considered larger coalitions up until now.

Nonetheless, the following definition and theorem show that, in the marriage market, this simplification does not make any difference, as nothing is lost by ignoring coalitions formed by more than two individuals.

Definition 3.16. The *core* of a marriage market is the set of undominated matchings.

We denote the core of the marriage market $(M, W; >)$ by $\mathcal{C}(P)$.

Theorem 3.17. *Let $(M, W; >)$ be a marriage market. Then $\mathcal{C}(P) = \mathcal{S}(P)$.*

Proof.

$\subseteq$) Let us see that $\forall \mu \in \mathcal{M}(M, W)$, $\mu \notin \mathcal{S}(P) \Rightarrow \mu \notin \mathcal{C}(P)$.
If μ is individually irrational, then it is dominated via a coalition formed by a single individual. On the other hand, if it is unstable via a blocking pair $\{m, w\}$ such that $m >_w \mu(w)$ and $w >_m \mu(m)$, then it is dominated via the coalition $\{m, w\}$.

$\supseteq$) Let us see that $\forall \mu \in \mathcal{M}(M, W)$, $\mu \notin \mathcal{C}(P) \Rightarrow \mu \notin \mathcal{S}(P)$.
If μ is not in the core, then it is dominated via some coalition A by some other matching ν. In the case that μ is individually rational, $\forall w \in A$, $\nu(w) >_w \mu(w)$ and so $\nu(w) \in M \cap A$.
Assume $w \in A$ and $m = \nu(w)$. Then $w >_m \mu(m)$, and μ is blocked by $\{m, w\}$.

□

We will now analyse the algebraic structure of the set of stable matchings. Let us consider the partial orders $\succeq_W$, $\succ_W$, $\succeq_M$ and $\succ_M$ restricted to $\mathcal{S}(P)$. These are partial orders, because there are stable matchings which might not possibly be compared in this way. In fact, $\mathcal{S}(P)$ is a distributive lattice under these partial orders. Let us recall the definition of lattice.

Definition 3.18. A *lattice* $(L, \geq)$ is a partially ordered set such that any two elements x and y have a supremum, denoted by $x \vee y$, and an infimum denoted by $x \wedge y$.

[5]By stating false preferences; we will see how this works in Section 3.4.

A lattice is *complete* if each of its subsets $X \subseteq L$ has a supremum and an infimum in L. A *distributive lattice* $(L, \geq)$ is a lattice such that, $\forall x, y, z \in L$,

(a) $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$

(b) $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$

Given a lattice $(L, \geq)$, its *dual lattice* is a lattice $(L, \geq^d)$ such that $\forall a, b \in L$, $b \geq^d a \iff a \geq b$.

The following theorem helps to understand better the collision and coalescence of interests among agents in these markets. The more general result is later proved for the college admissions problem.

Theorem 3.19. (Conway as cited by Knuth, 1976) *In the marriage problem $(M, W; >)$, if preferences are strict, $\mathcal{S}(P)$ is a distributive lattice under the partial order $\succeq_W$, dual to the partial order $\succeq_M$. If we consider the partial order $\succeq_W$ (resp. $\succeq_M$), then the maximum (resp. minimum) element of the lattice is μ_W and the minimum (resp. maximum) element is μ_M.*

That is, intuitively, every time we move closer to the W-optimal matching, we consequently must move further from the M-optimal matching. Actually, an immediate corollary of Theorem 3.19 is that for any two matchings $\mu, \nu \in \mathcal{S}(P)$, $\mu \succ_M \nu \iff \nu \succ_W \mu$.

Another interesting result regarding stable matchings is that the set of self-matched agents of a marriage market $(M, W; >)$ is the same in every stable matching. This result is called the Rural Hospital Theorem, and we will study it in the following chapter because it is also true for the college admissions problem.

3.4 Strategic behaviour

We have not yet dug into the different strategies players can adopt, which are the preferences agents can state. If agents are likely to say the truth regarding their preferences is a question that arises naturally after all the previous results. We will focus on the incentives they have for truth-telling under the deferred acceptance mechanism.

Of course, if agents had incentives to lie, then the algorithm would not be useful even if it did yield stable outcomes with respect to the stated preferences (because these would not be the true preferences).

Remark 3.20. On the one hand, when we consider the marriage *problem* $(M, W; >)$, we are fixing P, that is, we are assuming that either agents will not state any list differing from their true preferences or that the preferences they state are their true ones. On the other hand, when we consider the marriage *game* $(\mathcal{I}, S, u)$, we are taking for granted that agents may choose any element $s \in S$ as a strategy, and thus become strategic players who may potentially lie if they wish.

Therefore, these are two completely different approaches to the same market. Reason why, throughout this section, we will consider the same marriage market as $(M, W; >)$ or as $(\mathcal{I}, S, u)$ depending on which one fits best our purpose at every given moment. We will use P to refer to the preference profile with true preferences submitted.

Proposition 3.21. *Truth-telling is not a dominant strategy for all agents of the marriage game $(\mathcal{I}, S, u)$ when the deferred acceptance algorithm is used.*

Actually, we will prove the following theorem, which is even stronger.

Theorem 3.22. (Impossibility Theorem) *There is no mechanism for the marriage game $(\mathcal{I}, S, u)$ which yields a matching in $\mathcal{S}(P)$ and makes it an dominant strategy for all agents to state their true preferences.*

Proof. We only need to find a marriage game for which no stable matching mechanism makes truth-telling a dominant strategy. Consider the marriage game $(M \cup W, S, u)$, where $M = \{m_1, m_2, m_3\}$, $W = \{w_1, w_2, w_3\}$, and the (true) preference orders are the following

$$\begin{aligned}
>_{w_1}&: \; m_2 >_{w_1} m_1 >_{w_1} m_3 >_{w_1} w_1;\\
>_{w_2}&: \; m_1 >_{w_2} m_2 >_{w_2} m_3 >_{w_2} w_2;\\
>_{w_3}&: \; m_1 >_{w_3} m_2 >_{w_3} m_3 >_{w_3} w_3;\\
>_{m_1}&: \; w_1 >_{m_1} w_3 >_{m_1} w_2 >_{m_1} m_1;\\
>_{m_2}&: \; w_3 >_{m_2} w_1 >_{m_2} w_2 >_{m_2} m_2;\\
>_{m_3}&: \; w_1 >_{m_3} w_2 >_{m_3} w_3 >_{m_3} m_3.
\end{aligned}$$

If all agents submit the preference profile $P \in \mathcal{P}(M, W)$ according to their true preference orders, there are six individually rational matchings in which nobody remains single:

$$\mu_1 = \begin{pmatrix} w_1 & w_2 & w_3 \\ m_1 & m_2 & m_3 \end{pmatrix}, \quad \mu_3 = \begin{pmatrix} w_2 & w_1 & w_3 \\ m_1 & m_2 & m_3 \end{pmatrix}, \quad \mu_5 = \begin{pmatrix} w_2 & w_3 & w_1 \\ m_1 & m_2 & m_3 \end{pmatrix},$$
$$\mu_2 = \begin{pmatrix} w_1 & w_3 & w_2 \\ m_1 & m_2 & m_3 \end{pmatrix}, \quad \mu_4 = \begin{pmatrix} w_3 & w_1 & w_2 \\ m_1 & m_2 & m_3 \end{pmatrix}, \quad \mu_6 = \begin{pmatrix} w_3 & w_2 & w_1 \\ m_1 & m_2 & m_3 \end{pmatrix}.$$

Matchings μ_1 and μ_6 are blocked by the pair $\{m_2, w_1\}$, μ_3 is blocked by $\{m_1, w_3\}$, and μ_5 is blocked by the pair $\{m_1, w_1\}$; μ_2 and μ_4 have no blocking pairs. Thus, $\mathcal{S}(P) = \{\mu_M, \mu_W\}$, with $\mu_M = \mu_2$ and $\mu_W = \mu_4$.

Now, suppose w_1 lied and stated her following false preferences: $P'(w_1) = m_2, m_3, m_1$. Then, we would have a new marriage problem $(M, W; >')$, with $>'$ the preference order $(>'_x)_{x \in M \cup W}$ such that $>'_x \equiv >_x$, $\forall x \in M \cup W \backslash \{w_1\}$, and $>'_{w_1}$: $m_2 >'_{w_1} m_3 >'_{w_1} m_1 >'_{w_1} w_1$. The set $\mathcal{M}(M, W)$ would obviously remain unchanged, and the only matching not blocked by any couple would be $\mu_4 = \mu_W$. Therefore, we would have $\mathcal{S}(P') = \{\mu_W\}$.

Suppose instead that it was m_1 who lied by stating his following false preferences: $P''(m_1) = w_1, w_2, w_3$. Then, the new problem would be $(M, W; >'')$, with $>''$ the preference order $(>''_x)_{x \in M \cup W}$ such that $>'' \equiv >_x$, $\forall x \in (M \cup W) \backslash \{m_1\}$, and $>''_{m_1}$: $w_1 >_{m_1} w_2 >_{m_1} w_3 >_{m_1} m_1$. In this case, the only stable matching would be μ_M; that is, $\mathcal{S}(P'') = \{\mu_M\}$.

Let us now consider the only two possibilities regarding $f(P)$, assuming that f is a stable-matching yielding mechanism with reference to preferences P.

(i) $f(P) = \mu_M$. In this case, w_1 has incentives to misrepresent her preferences, because by stating $P'(w_1)$ instead of $P(w_1)$ she would obtain $f(P') = \mu_W$, and $\mu_W \succ_{w_1} \mu_M$.

(ii) $f(P) = \mu_W$. In this case, m_1 has incentives to misrepresent his preferences, because by stating $P''(m_1)$ instead of $P(m_1)$ he would obtain $f(P'') = \mu_M$, and $\mu_M \succ_{m_1} \mu_W$.

□

This is not a very optimistic result. However, our next theorem will help us see why, despite not being optimal, the DAA seems to be as good as possible.

Theorem 3.23. (Dubins and Freedman, 1981) *Let $(M, W; >)$ be a marriage problem. Then, when a mechanism that produces the W-optimal (resp. M-optimal) stable matching is used, it is a dominant strategy -maybe only weakly- for all $w \in W$ (resp. $m \in M$) to state their true preferences.*

Proof. See [9]. □

Obviously, the DAA falls into the category of mechanisms for which this theorem works. In addition, agents on the non-proposing side of the market do not have any incentive to misrepresent their first choice.[6]

The last result of this chapter will help us see the relation between the marriage problem and Chapter 2 more clearly.

Theorem 3.24. *Let $(\mathcal{I}, S, u)$ be a marriage game with $\mathcal{I} = M \cup W$ and $S = \mathcal{P}(M, W)$. If a procedure that yields the W-optimal stable matching is used and the strategy profile chosen, $s^* \in S$, is a Nash equilibrium with none of its strategies weakly dominated, then the resulting matching μ^* is stable with respect to the true preferences P; that is, $\mu^* \in \mathcal{S}(P)$.*

Proof. Let us consider the related marriage problem $(M, W; >)$, and a W-optimal stable matching procedure $f : \mathcal{P}(M, W) \to \mathcal{M}(M, W)$. Let us denote $s^* = (s^*_{m_1}, \ldots, s^*_{m_p}, s^*_{w_1}, \ldots, s^*_{w_n})$, and let us assume that $\mu^* = f(s^*) \notin \mathcal{S}(P)$. Moreover, we can assert without loss of generality that this procedure is the DAA with women proposing.

Therefore, $\exists m_i, w_j$ such that $w_j >_{m_i} \mu^*(m_i)$ and $m_i >_{w_j} \mu^*(w_j)$. By Theorem 3.23, $s^*_{w_j}$ must be w_j's true preferences; but then, w_j must have proposed to m_i before proposing to $\mu^*(w_j)$, and m_i must have declined this proposal in favour of a more preferred one.

Let us consider the preference profile $s' \in \mathcal{P}(M, W)$, with $s'_x = s^*_x \; \forall x \in (M \cup W) \backslash \{m_i\}$ and s'_{m_i} the strategy of man m_i which ranks w_j as his top choice and leaves everything else unchanged (with respect to $s^*_{m_i}$). Clearly, if $\mu' = f(s')$, then $\mu'(m_i) = w_j$, since w_j proposes to m_i at the same step than before but she is not rejected.

Hence, it is clear that regardless of which particular utility function u is, $u_{m_i}(s^*) < u_{m_i}(s^*_{-m_i}, s'_{m_i})$, which contradicts the fact that s^* is a Nash equilibrium. □

Obviously, the analogous result swapping men and women is also true.

To conclude, there is a question that arises naturally after the study of this chapter, and is that of how many outcomes actually form the set $\mathcal{S}(P)$. There is a lot of bibliography regarding this issue, which will not be studied here, but the main conclusion is that finding this number is NP-hard.[7]

[6] See [23].
[7] An asymptotic bound from above is found in [8].

Chapter 4

The college admissions problem

We turn now to study the college admissions problem. In this problem, there are two distinct sets of agents to be matched, but unlike in the marriage problem, agents from one of the two sets might end up assigned to many agents on the other side. This is why this kind of problem is also called many-to-one matching problems.[1]

There are multiple models in real life that belong to this category, such as labour markets in which firms employ many workers, or universities that enrol many students, while workers usually work for a single firm and students hardly attend more than one college at a time. In these models, one side of the market is usually made of individuals and the other one of institutions.

When comparing many-to-one matching to the previously seen one-to-one matching, we will observe that there are differences, but also noteworthy similarities. Many of the differences come from the fact that institutions need to compare subsets of agents from the other side, generally containing more than one individual. It was initially believed that the college admissions problem could be thought of as a marriage problem, where each one of the seats available at a given college would be treated as a single individual.[2] We will see why this is not the case, even though this view will prove useful for some of the results.

The model

Let us start to build the formal model: there are two disjoint and finite sets, students $S = \{s_1, \ldots, s_m\}$, and colleges $C = \{c_1, \ldots, c_n\}$, and the aim of the market is to match them. Each student can attend one college, and each college $c_i \in C$ can admit at most its *quota* of students, $q_i \in \mathbb{N}$; that is, the quota of college c is the number of positions or seats it has to offer. We define the quota vector $q := (q_1, \ldots, q_n)$.

Similarly to the marriage problem, every student $s \in S$ is endowed with preferences over $C \cup \{s\}$ that induce a strict order relation $>_s$. Student s submits to the central administration, which will try to find a solution to the problem, an ordered list $P(s)$ over $C \cup \{s\}$

[1] The natural generalisation, many-to-many matching problems, is interestingly studied in [3].

[2] In [9] and [23] this is taken for granted, and in [15] there is a (mistaken) proof of the equivalence between these two problems.

according to her order relation $>_s$ (or not, as we have seen for the marriage market; for now, let us suppose she does).[3]

Analogously, every college $c \in C$ has preferences over $S \cup \{c\}$ that induce a strict order relation $>_c$. The ordered list college c submits to the central administration will be denoted by $P(c)$.

We will obviously assume that all elements of $>:= (>_x)_{x \in S \cup C}$ are complete, transitive and strict order relations. The kind of notation is the same as the one used in the marriage problem, so we will shorten the preference lists by leaving only the acceptable options.[4]

A preference profile is denoted by $P = P(s_1) \times \ldots \times P(s_m) \times P(c_1) \times \ldots \times P(c_n)$, where $P(x)$ is the preference list submitted by $x \in S \cup C$ in accordance to the order $>_x$. The set of all possible preference profiles is denoted by $\mathcal{P}(S, C)$.

To summarize, a college admissions problem is denoted by the tetrad $(S, C, q; >)$. If we consider it as a game $G = (\mathcal{I}, \Sigma, u)$, the elements are

(a) $\mathcal{I} = S \cup C$,

(b) $\Sigma = \mathcal{P}(S, C)$, and

(c) $\forall x \in \mathcal{I}$, a utility function $u_x : \Sigma \to \mathbb{R}$.[5]

The rules of the game $(\mathcal{I}, \Sigma, u)$ are the following:

(a) any pair consisting of one student and one college that agrees to be matched may do so,

(b) any college might decide to leave some sit empty, that is, it may prefer to leave its quota unfilled rather than to admit some of the students in S,

(c) any student may remain unmatched if she wishes, and

(d) no college $c \in C$ may admit more than its quota q_c of students and no student can enrol more than one college.

An outcome of the college admissions problem is an assignment of students to colleges (and vice-versa), such that each student is matched to at most one college and each college is matched at most to its quota of students. The convention is that if a student is not matched to any college at a given matching, then she is self-matched, and if a college does not fill its quota, then it is matched to itself in each of its empty positions.

We will call, for a given set A, an *unordered family of elements of A* a collection of elements of A, not necessarily distinct, in which the order is irrelevant.[6] We can now proceed to define a matching in this model.

Definition 4.1. Let $(S, C, q; >)$ be a college admissions problem. A *matching* μ is a map from $S \cup C$ into the set of unordered families of elements of $S \cup C$ such that

[3] Therefore, for the moment we will use indistinctly $>_s$ and $P(s)$. The same will be done for colleges.

[4] In the preference list of agent $x \in S \cup C$, there will only be those elements $y \in (S \cup C) \backslash \{x\}$ such that $y >_x x$.

[5] Always keeping in mind that u_x is actually a map over the set of outcomes, but because it is equivalent to a map over Σ, there is no loss of generalisation by considering its domain is Σ.

[6] That is, the difference between an unordered family of elements of A and a subset of A is that in the former, a given element of A may appear more than once.

(a) $|\mu(s)| = 1$, and $\mu(s) \in C \cup \{s\}$ $\forall s \in S$,

(b) $|\mu(c)| \leq q_c$ $\forall c \in C$, and if $r = |\mu(c)| < q_c$, then $\mu(c)$ contains $q_c - r$ copies of c, and

(c) $\mu(s) = c \iff s \in \mu(c)$.

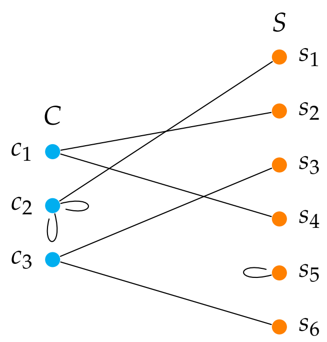
The set of all possible matchings of the college admissions problem $(S, C, q; >)$ is denoted by $\mathcal{M}(S, C)$. Recalling the marriage problem, we will represent matchings in the following way:

Example 4.2. Let $(S, C, q; >)$ be a college admissions problem, where $S = \{s_1, \ldots, s_6\}$, $C = \{c_1, c_2, c_3\}$, and $q = (2, 3, 2)$. In the matching

$$\mu = \begin{pmatrix} c_1 & c_2 & c_3 & (s_5) \\ s_2\ s_4 & s_1\ c_2\ c_2 & s_3\ s_6 & s_5 \end{pmatrix},$$

the assignment of college c_1, which has quota 2, is $\{s_2, s_4\}$; similarly, the assignment of college c_3 is $\{s_3, s_6\}$. On the other hand, college c_2 does not fill its quota $q_2 = 3$, and enrols only student s_1, while student s_5 remains unmatched.

Just as in one-to-one matchings, we can represent μ by means of a graph as follows.



Notice that q_i is the number of edges attached to node c_i, $\forall i = 1, 2, 3$.

4.1 Preferences and stability

As we have briefly mentioned in the beginning of this chapter, the preferences described above are not enough in this model to compare matchings, since colleges with quota greater than one are unable to compare different matchings in terms of preferences over single students. That is, they will have to compare groups of students, which is impossible with the tools we have so far developed. Thus, we need to consider, for every college $c \in C$, an order relation over all the assignments that can arise from a matching; this order will be denoted by $>^{\#}_c$.

Definition 4.3. Consider the college admissions problem $(S, C, q; >)$. Let $>_c$ be the ordered preferences of college c over $S \cup \{c\}$, and let $>^{\#}_c$ be a preference relation of c over all sets of students (union itself) it could be assigned at any matching $\mu \in \mathcal{M}(S, C)$. Then, $>^{\#}_c$ is *responsive* to $>_c$ if, $\forall \mu' \in \mathcal{M}(S, C)$ such that $\mu'(c) = (\mu(c) \cup \{s_i\}) \backslash \{\sigma\}$, where $\sigma \in \mu(c)$ and $s_i \notin \mu(c)$, then

$$\mu'(c) >_c^{\#} \mu(c) \iff s_i >_c \sigma.$$

In other words, preferences $>_c^{\#}$ are responsive to $>_c$ if, when faced with two assignments differing in only one element, c prefers the assignment with its preferred element. This is a natural extension of colleges' preferences over individuals, and so we will henceforth assume that colleges have responsive preferences, on top of they being complete and transitive (and strict over individuals).

Notice that there may exist many different responsive preference orderings $>_c^{\#}$ for every $>_c$.

Example 4.4. Consider the college admissions problem $(S, C, q; >)$ with $S = \{s_1, \ldots, s_4\}$. Let $c \in C$ be a college with quota $q_c = 2$ and preferences $P(c)$ over $\{s_1, s_2, s_3, s_4, c\}$ given by $P(c) = s_1, s_2, s_3, s_4$. Two possible responsive preferences over entering classes are

$$P_1^{\#}(c) = \{s_1, s_2\}, \{s_1, s_3\}, \{s_1, s_4\}, \{s_2, s_3\}, \{s_2, s_4\}, \{s_3, s_4\}, \{s_1, c\}, \{s_2, c\}, \{s_3, c\}, \{s_4, c\},$$

and

$$P_2^{\#}(c) = \{s_1, s_2\}, \{s_1, s_3\}, \{s_2, s_3\}, \{s_1, s_4\}, \{s_2, s_4\}, \{s_3, s_4\}, \{s_1, c\}, \{s_2, c\}, \{s_3, c\}, \{s_4, c\}.$$

Stability

The notions regarding stability are analogous to those in the marriage problem.

Definition 4.5.

- A matching $\mu \in \mathcal{M}(S, C)$ is *individually irrational* if $\exists s \in S, c \in C$ such that $\mu(s) = c$ and either s is unacceptable to c, or c is unacceptable to s. That is, if $c >_c s$ or $s >_s c$.
- A matching $\mu \in \mathcal{M}(S, C)$ is *blocked by the college-student pair* $\{c, s\}$, with $c \in C$, $s \in S$, if $\mu(s) \neq c$, $c >_s \mu(s)$, and $\exists \sigma \in \mu(c)$ such that $s >_c \sigma$.
- A matching $\mu \in \mathcal{M}(S, C)$ is *stable* if it is not blocked by any individual agent or college-student pair. $\mu \in \mathcal{M}(S, C)$ is *unstable* if it is not stable.

We will denote the set of all stable matchings by $\mathcal{S}(P)$.

At first sight, it seems clear that this definition of stability might not prove useful when trying to find a solution for the college admissions problem, since this model involves preferences over sets of agents, and the coalitions that could potentially conspire to improve their results are much more heterogeneous and diverse than those in the marriage market. We will hence shed light on this matter and see that when preferences are responsive, nothing is lost by considering only simple college-student blocking pairs.

Definition 4.6. Given the college admissions problem $(S, C, q; >)$, a matching $\mu \in \mathcal{M}(S, C)$ is *group unstable* if there exists another matching $\nu \in \mathcal{M}(S, C)$ and a coalition $A \subset S \cup C$ such that $\forall s \in A \cap S$ and $\forall c \in A \cap C$,

(a) $\nu(s) \in A \backslash \{s\}$,

(b) $\nu(s) >_s \mu(s)$ (if μ is individually rational, (b) $\Longrightarrow$ (a)),

(c) $\nu(c) >_c \mu(c)$, and

(d) $\sigma \in \nu(c) \implies \sigma \in A \cup \mu(c)$, i.e., every student in $\nu(c)$ is either a new student from A or one of c's old students in $\mu(c)$.

In this case, it is also said that μ is *blocked by the coalition A*.

In view of this, if $\mu \in \mathcal{M}(S, C)$ is blocked by the coalition of colleges and students A, then they could all get a better assignment by matching among themselves (notice that if μ is individually rational, A must have agents from both C and S). If such coalition does not exist, then the matching μ is called *group stable*. Surprisingly, this notion is equivalent to that of stability we have seen before, as stated in the following lemma.

Lemma 4.7. *Given the college admissions problem $(S, C, q; >)$, a matching $\mu \in \mathcal{M}(S, C)$ is group stable $\iff \mu \in \mathcal{S}(P)$.*

Proof.

$\Rightarrow$) If $\mu \notin \mathcal{S}(P)$ then there is a single agent, or a student-college pair that blocks μ, namely a coalition formed by a single agent or by a pair of agents that blocks μ; therefore, μ is group unstable.

$\Leftarrow$) If μ is blocked by a coalition $A \subset S \cup C$ that can achieve a matching $\nu \in \mathcal{M}(S, C)$ every member of A prefers to μ, then for $c \in A$, $\nu(c) >_c \mu(c)$, and hence $\exists s, \sigma$ such that $s \in \nu(c) \backslash \mu(c)$, $\sigma \in \mu(c) \backslash \nu(c)$ and $s >_c \sigma$. Therefore, $s \in A$, and $c >_s \mu(s)$, so the matching μ is unstable via the pair $\{s, c\}$.

□

Thus, there are two immediate simplifications we can implement after this lemma. The first one is that we can concentrate on small coalitions of a single agent or a student-college pair when searching for stable matchings. The second one is that when preferences are responsive, which we are assuming is true, we can find stable matchings by using only the preferences P over individuals, and so the responsive preferences of college c over groups of students, $P^{\#}(c)$, are not needed anymore. Knowing this, it would not be a surprise to see that many of the results seen for the marriage problem carry over to the college admissions problem.

4.2 The related marriage problem

In this section, we are going to consider a particular college admissions problem and slightly adjust it in a way that it becomes a marriage problem. The key aspect to take into account is that the marriage problem is a particular case of a college admissions problem in which all colleges have quota one. In this section, responsive preferences of colleges over sets of students are not taken for granted.

Let $C = \{c_1, \ldots, c_n\}$ be the set of colleges and $S = \{s_1, \ldots, s_m\}$ the set of students; the quota of college c_k is $q_k \in \mathbb{N}$, and the preference profile $P = P(c_1) \times \ldots \times P(c_n) \times$

$P(s_1) \times \ldots \times P(s_m) \in \mathcal{P}(S,C)$. If we replace every college $c_k \in C$ with q_k copies of itself, i.e, with the string $c_k^1, c_k^2, \ldots, c_k^{q_k}$, we can assume that each one of these copies is a college with quota one; this is why, in this section, we only need to consider preferences over individuals and not over greater sets. Moreover, the preferences over individuals of each one of these copies are exactly those of c_k.

One problem that may arise is the fact that students do not have strict preferences over these copies of c_k, and strict preferences are a key limitation in our model; however, this is easily fixed by replacing c_k in the preference list of every student s with the string $c_k^1, \ldots, c_k^{q_k}$, in that order. This way, we can impose $c_k^i >_s c_k^j, \forall i < j, \forall s \in S$.

Thus, we have, for the college admissions problem $(S, C, q; >)$, a related marriage problem $(S, \widetilde{C}; \widetilde{>})$, where $\widetilde{C} = \{c_1^1, \ldots, c_1^{q_1}, \ldots, c_n^1, \ldots, c_n^{q_n}\}$, and $\widetilde{>} = (\widetilde{>}_x)_{x \in S \cup \widetilde{C}}$ are the preference orders corresponding to $S \cup \widetilde{C}$. The preference profile of this marriage problem is denoted by $\widetilde{P}$.

The correspondence between matchings in the original college admissions problem and the related marriage market we have built is clear now; moreover, the stability of a matching is preserved in the transition from a college admissions problem to its related marriage problem. Regarding this issue, we have the following lemma.

Lemma 4.8. *Given the college admissions problem $(S, C, q; >)$, a matching $\mu \in \mathcal{M}(S,C)$ is stable if, and only if, the corresponding matching of the related marriage problem $(S, \widetilde{C}; \widetilde{>})$, $\widetilde{\mu} \in \mathcal{M}(S, \widetilde{C})$, is stable.*

This lemma enables us to generalise some of the results of the marriage problem to the college admissions problem, such as the fact that $\mathcal{S}(P) \neq \emptyset$ for every marriage problem.

Nonetheless, not all of the results do generalize, and there are remarkable differences. These differences arise when we study the strategic decisions available to the agents or the structure of the set of stable matchings. For example, we cannot conclude from Lemma 4.8 anything regarding the preferences of colleges over different matchings, because now stable matchings can be identified using only the preferences of colleges over single students.

4.3 The deferred acceptance algorithm

Let us consider the deferred acceptance algorithm in the case of the college admissions problem $(S, C, q; >)$. In particular, we will display the deferred acceptance algorithm with students proposing. The college-proposing deferred acceptance algorithm is obtained by transposing the roles of colleges and students, but taking into account that on each step, colleges make as many proposals as empty seats they have (unless there are not enough students to make that many proposals).

Step 1. (a) Each student $s \in S$ proposes to her top choice among those colleges for which she is acceptable.

(b) Each college $c \in C$ rejects all but the best q_c students among those students who proposed to it, where q_c denotes the quota of c. Those that remain are tentatively assigned a seat at c.

⋮

Step k. (a) Each student who has been rejected in the step $k-1$ proposes to her top choice among those colleges that have not yet rejected her and for which she is acceptable, if there exists such college.

(b) Each college c rejects all but the best q_c students among those who have just proposed and those who were tentatively assigned to it at the step $k-1$. The ones that are not rejected are tentatively assigned one seat at c.

The algorithm stops in the step in which no proposal is rejected or there are no more possible proposals to make, and the tentative assignments at that moment become definitive; if a student has been rejected by all the colleges to which she has proposed, then she is not assigned to any college and remains self-matched. Similarly, if a college has not filled its quota by the step the algorithm stops, then it remains self-matched in all of its empty positions.

Notice that there is a difference between the DAA for the marriage problem and the one we have just described for the college admissions problem: in the latter, at every step, no university receives proposals from unacceptable students. As stable matchings are always individually rational, this can be done without loss of generality,[7] and the algorithm will probably stop earlier than otherwise.

Example 4.9. Consider the college admissions problem $(S, C, q; >)$, where $C = \{c_1, c_2, c_3\}$, $S = \{s_1, s_2, s_3, s_4, s_5\}$, $q = (2, 2, 1)$, and the submitted preferences are given by

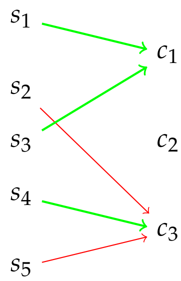
$$
\begin{array}{ll}
P(s_1) = c_1, c_3, c_2; & P(c_1) = s_2, s_3, s_1, s_5; \\
P(s_2) = c_3, c_2, c_1; & P(c_2) = s_4, s_5, s_2; \\
P(s_3) = c_1, c_2, c_3; & P(c_3) = s_3, s_4, s_1, s_5, s_2. \\
P(s_4) = c_1, c_3, c_2; & \\
P(s_5) = c_3, c_1, c_2. &
\end{array}
$$

Step 1. Every student proposes to her first-choice college but s_4, who has to apply to her second choice because she is unacceptable to her first choice.

- c_1 receives proposals from s_1 and s_3. It keeps both tentatively assigned.
- c_2 receives no proposals, and so it remains self-matched in all of its positions.
- c_3 receives proposals from s_2, s_4 and s_5. As $s_4 >_{c_3} s_5 >_{c_3} s_2$, it rejects s_2 and s_5, and keeps s_4 tentatively assigned.

Step 1 can be represented by means of the following directed graph, where green edges represent accepted proposals and red ones represent declined proposals:

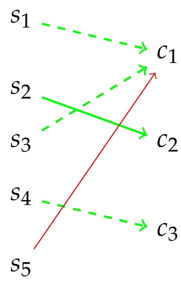
[7] Because we have removed only those proposals that would be rejected for certain.



Step 2. Only s_2, s_5 make proposals in this step: s_2 proposes to c_2, while s_5 proposes to c_1.

- $s_3 >_{c_1} s_1 >_{c_1} s_5$, so c_1 rejects the received proposal and does not change its tentative assignment.
- c_2 keeps s_2 tentatively assigned.
- c_3 receives no new proposals, so its tentative assignment remains unchanged.

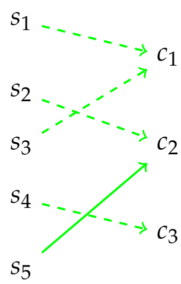
Step 2 is represented below. Dashed edges indicate tentative assignments from the previous step, while the colour of the edge indicates if that proposal has been accepted (green) or rejected (red).



Step 3. s_5 is the only student to make a proposal in this step, and proposes to c_2.

- c_1 receives no proposals, so its tentative assignment remains unchanged.
- c_2 accepts s_5's proposal, and so its tentative assignment is $\{s_2, s_5\}$.
- c_3 receives no new proposals, so its tentative assignment remains unchanged.

Step 3 is represented below in the same way as be have represented step 2.



In Step 3 no proposal has been rejected, so the algorithm stops and the resulting matching for this college admission problem using the DAA with students proposing is

$$\mu = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1\, s_3 & s_2\, s_5 & s_4 \end{pmatrix}.$$

4.4 Strategies and truth-telling

Just as in the marriage market, when preferences are strict, there is a C-optimal stable matching that every college likes at least as well as any other stable matching, μ_C, and a S-optimal stable matching that every student likes at least as well as any other stable matching, μ_S. The following two definitions carry over immediately from the marriage problem.

Definition 4.10. Consider the college admissions problem $(S, C, q; >)$. For every element $x \in S \cup C$, we can define two order relations over the set $\mathcal{M}(S, C)$: $\forall \mu, \nu \in \mathcal{M}(S, C)$, we write $\mu \succeq_x \nu$ to denote that x likes μ at least as well as ν (that is, either $\mu(x) >_x \nu(x)$ or $\mu(x) = \nu(x)$), and $\mu \succ_x \nu$ to denote that x strictly prefers μ to ν ($\Leftrightarrow \mu(x) >_x \nu(x)$).

We can also define order relations over $\mathcal{M}(S, C)$ for the sets S and C. We write $\mu \succeq_S \nu$ to indicate that $\forall s \in S$, $\mu \succeq_s \nu$ and $\exists s \in S$ such that $\mu \succ_s \nu$, and $\mu \succ_S \nu$ to indicate that $\forall s \in S$, $\mu \succ_s \nu$. Orders $\succeq_C$ and $\succ_C$ are defined analogously.

Definition 4.11. Given the college admissions problem $(S, C, q; >)$, $\mu \in \mathcal{S}(P)$ is *S-optimal* if $\forall \nu \in \mathcal{S}(P)$, $\mu \succeq_S \nu$. A *C-optimal* matching is defined analogously.

Just as in the marriage problem, the S-optimal stable matching is unique, and is the one obtained when using the deferred acceptance algorithm with students proposing, and similarly the C-optimal stable matching is the one obtained when colleges take the initiative. The following theorem, however, is not equivalent to that in the marriage problem.

Theorem 4.12. *Given the strict preference profile $P \in \mathcal{P}(S, C)$, the S-optimal matching of the college admissions problem $(S, C, q; >)$ is weakly Pareto optimal for the students, but the C-optimal stable matching might not be even weakly Pareto optimal for the colleges.*

Proof. On the one hand, for the part of the proof regarding students the result carries over immediately from the marriage problem. On the other hand, the result concerning colleges will we shown by means of a counterexample. We will find a college admissions problem $(S, C, q; >)$ so that every college in C prefers another matching to the C-optimal matching.

Consider $(S, C, q; >)$ with $S = \{s_1, s_2, s_3, s_4\}$, $C = \{c_1, c_2, c_3\}$ and $q = (1, 1, 2)$. Let the preference profile P be given by

$$\begin{array}{ll} P(s_1) = c_3, c_1, c_2; & P(c_1) = s_3, s_2, s_4, s_1; \\ P(s_2) = c_1, c_3, c_2; & P(c_2) = s_4, s_3, s_2, s_1; \\ P(s_3) = c_2, c_3, c_1; & P(c_3) = s_3, s_2, s_4, s_1. \\ P(s_4) = c_3, c_2, c_1. & \end{array}$$

There are twelve individually rational matchings in which no agent is self-matched, and only one of them is stable:

$$\mu_1 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1 & s_2 & s_3\, s_4 \end{pmatrix}, \quad \mu_5 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2 & s_3 & s_1\, s_4 \end{pmatrix}, \quad \mu_9 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_3 & s_4 & s_1\, s_2 \end{pmatrix},$$
$$\mu_2 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1 & s_3 & s_2\, s_4 \end{pmatrix}, \quad \mu_6 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2 & s_4 & s_1\, s_3 \end{pmatrix}, \quad \mu_{10} = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_4 & s_1 & s_2\, s_3 \end{pmatrix},$$
$$\mu_3 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_1 & s_4 & s_2\, s_3 \end{pmatrix}, \quad \mu_7 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_3 & s_1 & s_2\, s_4 \end{pmatrix}, \quad \mu_{11} = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_4 & s_2 & s_1\, s_3 \end{pmatrix},$$
$$\mu_4 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2 & s_1 & s_3\, s_4 \end{pmatrix}, \quad \mu_8 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_3 & s_2 & s_1\, s_4 \end{pmatrix}, \quad \mu_{12} = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_4 & s_3 & s_1\, s_2 \end{pmatrix}.$$

Matchings $\mu_1, \mu_2, \mu_3, \mu_{10}, \mu_{11}$ and μ_{12} are blocked by the pair $\{c_1, s_2\}$, μ_4 and μ_7 by the pair $\{c_2, s_3\}$, and μ_6, μ_8, μ_9 are blocked by the pairs $\{c_3, s_4\}, \{c_3, s_2\}$ and $\{c_3, s_3\}$, respectively. Therefore, $\mathcal{S}(P) = \{\mu_5\}$, and $\mu_C = \mu_S = \mu_5$.

However, $\mu_9 \succ_C \mu_5$. That is, even though μ_5 is the C-optimal stable matching, there is another assignment, μ_9, that every college strictly prefers to μ_5. □

Therefore, we have seen that there may exist matchings that all colleges strictly prefer to the C-optimal stable outcome. Nonetheless, if there is a matching $\mu \in \mathcal{M}(S, C)$ such that at least one student is self-matched or one college does not fill its quota, the following theorem states that they could not have hoped for a better result.

Theorem 4.13. (Rural Hospital Theorem) *Given the problem $(S, C, q; >)$ with strict preferences over individuals $>$, the set of admitted students and filled seats is the same $\forall \mu \in \mathcal{S}(P)$. Moreover, any college $c \in C$ such that $|\mu(c)| < q_c$ for some $\mu \in \mathcal{S}(P)$, is assigned exactly the same set of students at every stable matching, that is, $\nu(c) = \mu(c)$, $\forall \nu \in \mathcal{S}(P)$.*

The motivation for this theorem is that when trying to assign hospitals to residents in the United States, it was observed that many hospitals in rural areas did not fill their quotas, and so the central administration tried to find a way to reassign these residents and solve this problem. Theorem 4.13 asserts that this is an impossible task, unless they let an unstable matching be the final outcome.

Let us now turn to the kind of questions agents are faced with when playing this game. The strategies of both students and colleges consist of stating lists of preferences. Our next theorem gives some insight into the kind of strategies colleges are likely to adopt. As seen in Theorem 4.12, the C-optimal stable matching might not be weakly Pareto optimal for the colleges, and so this theorem does not come as a surprise.

Theorem 4.14. *Given the college admissions problem $(S, C, q; >)$, there does not exist any stable matching procedure that makes it a dominant strategy for every college to state its true preferences.*

Proof. Consider the example used in the proof of Theorem 4.12. If there existed a procedure that always yielded a matching in $\mathcal{S}(P)$ and made it a dominant strategy for every college to say the truth, it would do so for any problem $(S, C, q; >)$, and so it is sufficient to show that it does not exist for this particular case.

Let us consider the problem as a game $G = (\mathcal{I}, \Sigma, u)$, with $\mathcal{I} = \{s_1, s_2, s_3, s_4, c_1, c_2, c_3\}$ and $\Sigma = \mathcal{P}(S, C)$. Recall that $\mathcal{S}(P) = \{\mu_5\}$, with

$$\mu_5 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_2 & s_3 & s_1\, s_4 \end{pmatrix},$$

and so any stable matching procedure must yield μ_5. Notice that even though μ_5 is the C-optimal matching, college c_3 is quite unhappy, because he has been given his third and fourth choices. Imagine that c_3 had lied and instead of submitting its true preferences $P(c_3) = s_3, s_2, s_4, s_1$, it had submitted the shortened preference list $P'(c_3) = s_2, s_1$. That is, let us suppose that the strategy profile chosen was $P' = P(s_1) \times \ldots \times P(s_4) \times P(c_1) \times P(c_2) \times P'(c_3)$.

In this case, every matching would be individually irrational, with the exception of

$$\mu_9 = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_3 & s_4 & s_1\, s_2 \end{pmatrix} \text{ and } \mu_{12} = \begin{pmatrix} c_1 & c_2 & c_3 \\ s_4 & s_3 & s_1\, s_2 \end{pmatrix},$$

and μ_{12} would still be blocked by the pair $\{c_1, s_2\}$. Therefore, $\mathcal{S}(P') = \{\mu_9\}$.

Now, recalling that μ_5 is the only stable matching of the problem $(S, C, q; >)$, it is clear that truth-telling is a dominated strategy for college c_3, because it prefers its assignment when stating $P'(c_3)$. □

Therefore, colleges (assuming they have the necessary information) in the college admission problem are likely to misrepresent their preferences in a way that they obtain a better result than at any stable matching. Clearly, if a single college can do so, a coalition of agents including at least one college can too.

However, students do not have this privilege for themselves, and their situation is as in the marriage market, because unlike colleges, their preferences in the college admissions problem are symmetrical to those of women or men in the marriage market.

Theorem 4.15. *Given the college admissions problem $(S, C, q; >)$, a stable matching mechanism that yields the S-optimal stable matching makes it a dominant strategy for all students to state their true preferences.*

Proof. If the theorem were false, there would be a student for whom stating her true preferences would not be a dominant strategy. Then, she would be able to benefit as well in the related marriage market $(S, \widetilde{C}; \widetilde{>})$. This contradicts Theorem 3.23. □

The positive side of this result is that students in a college admissions problem need not waste their time playing strategically: by knowing that their best option is to say the truth, they can submit their true preferences without fear of having made a mistake.

4.5 The set $\mathcal{S}(P)$

We will now study the set of stable matchings of the college admissions problem $(S, C, q; >)$. We assume throughout that preferences over individuals are strict, but colleges might be indifferent among various sets of students. In this section, the preference orders given by all agents respect their true orders.

In order to ease the notation, given an agent $x \in S \cup C$, her preference relation $>_x$ and two matchings $\mu, \nu \in \mathcal{M}(S,C)$, we will write $\nu(x) \geq_x \mu(x)$ to say that either $\nu(x) >_x \mu(x)$ or $\nu(x) = \mu(x)$.

Lemma 4.16. *Consider the college admissions problem $(S, C, q; >)$. Let $\mu, \nu \in \mathcal{S}(P)$, and suppose there $\exists c \in C$ such that $\mu(c) \neq \nu(c)$. Let $\widetilde{\mu}, \widetilde{\nu}$ be their corresponding stable matchings in the related marriage market $(S, \widetilde{C}; \widetilde{>})$. Then,*

$$\exists c^i \text{ position of } c \text{ such that } \widetilde{\mu}(c^i) \mathrel{\widetilde{>}_{c^i}} \widetilde{\nu}(c^i) \Rightarrow \forall c^j, \widetilde{\mu}(c^j) \mathrel{\widetilde{\geq}_{c^j}} \widetilde{\nu}(c^j),$$

with $i, j \leq q_c$.

In other words, this lemma says that there cannot be two matchings in $\mathcal{S}(P)$ such that there is a college indifferent between its two corresponding entering classes, unless this two classes are exactly the same.

Thus, when preferences over individuals are strict, colleges also have strict preferences over their entering classes when considering only the outcomes in $\mathcal{S}(P)$, and we can now assert that some more results from the marriage market carry over to the college admissions problem. The following theorem is one of them.[8]

Theorem 4.17. *Consider the college admissions problem $(S, C, q; >)$, and let $\mu, \nu \in \mathcal{S}(P)$. Then $\mu \succeq_C \nu \iff \nu \succeq_S \mu$.*

Proof. If we consider the related marriage market, $(S, \widetilde{C}; \widetilde{>})$, then

$$\mu \succeq_C \nu \iff \widetilde{\mu}(c_i) \mathrel{\widetilde{\geq}_{c_i}} \widetilde{\nu}(c_i)\ \forall c_i \in \widetilde{C}, \text{ and } \exists c_j \in \widetilde{C} \text{ such that } \widetilde{\mu}(c_j) \mathrel{\widetilde{>}_{c_j}} \widetilde{\nu}(c_j), \tag{4.1}$$

where $(\Rightarrow)$ follows from Lemma 4.16, while $(\Leftarrow)$ uses the responsiveness of c_i's preferences. Now, by definition, (4.1) $\iff \widetilde{\mu} \succeq_{\widetilde{C}} \widetilde{\nu} \iff \widetilde{\nu} \succeq_S \widetilde{\mu}$, by Theorem 3.19, and clearly $\widetilde{\nu} \succeq_S \widetilde{\mu} \iff \nu \succeq_S \mu$. $\square$

The following corollary is an immediate conclusion.

Corollary 4.18. *The optimal stable matching for one side of the market $(S, C, q; >)$ is the worst stable matching for the other side.*

These are quite remarkable results, since at first sight it would seem obvious that agents on one side of the market compete with each other for the best matching on the other side, but these two last results assert that when we narrow the possibilities down to stable matchings solely, all colleges prefer the C-optimal stable outcome and all students prefer the S-optimal matching, and what is more, μ_C is the less desirable stable matching for every student and μ_S is the less desirable stable matching for every college.

[8] It is the generalisation of the corollary of Theorem 3.19.

This resembles what we have previously seen regarding lattices in the marriage problem. Let us then turn again to lattices. Firstly, given two matchings of the college admissions problem $(S, C, q; >)$, $\mu, \nu \in \mathcal{M}(S, C)$, we define the maps $\lambda_{\mu,\nu}$ and $\eta_{\mu,\nu}$ from $S \cup C$ to the set of unordered families of elements of $S \cup C$ as follows:

$$\forall c \in C, \lambda_{\mu,\nu}(c) = \begin{cases} \mu(c) & \text{if } \mu(c) >_c \nu(c), \\ \nu(c) & \text{otherwise.} \end{cases}$$

$$\forall s \in S, \lambda_{\mu,\nu}(s) = \begin{cases} \mu(s) & \text{if } \nu(s) >_s \mu(s), \\ \nu(s) & \text{otherwise.} \end{cases}$$

Therefore, λ assigns each student the alternative she likes the least, and each college the one it likes the most. Below, we define η as the opposite of λ.

$$\eta_{\mu,\nu}(c) = \begin{cases} \mu(c) & \text{if } \nu(c) >_c \mu(c), \\ \nu(c) & \text{otherwise.} \end{cases}$$

$$\eta_{\mu,\nu}(s) = \begin{cases} \mu(s) & \text{if } \mu(s) >_s \nu(s), \\ \nu(s) & \text{otherwise.} \end{cases}$$

Notice that $\lambda_{\mu,\nu}$ and $\eta_{\mu,\nu}$ need not be matchings. For example, $\lambda_{\mu,\nu}$ will only be a matching in the case that $\lambda_{\mu,\nu}(s) = c \iff s \in \lambda_{\mu,\nu}(c)$. What is clear is that, when they are matchings, $\lambda_{\mu,\nu}$ (resp. $\eta_{\mu,\nu}$) is the least upper bound for $\{\mu, \nu\}$ under $\succeq_C$ (resp. $\succeq_S$) and the greatest lower bound under $\succeq_S$ (resp. $\succeq_C$). That is, when $\lambda_{\mu,\nu}, \eta_{\mu,\nu} \in \mathcal{M}(S, C)$,

$$\begin{aligned} \lambda_{\mu,\nu} &= \mu \vee_C \nu = \mu \wedge_S \nu, \\ \eta_{\mu,\nu} &= \mu \vee_S \nu = \mu \wedge_C \nu. \end{aligned}$$

We will now see that $\mathcal{S}(P)$ is a lattice under the partial orders $\succeq_C$ and $\succeq_S$. But first, we need to show that if μ and ν are stable matchings, then $\lambda_{\mu,\nu}$ and $\eta_{\mu,\nu}$ are stable matchings too. In order to ease the notation, we will henceforth get rid of the subscripts μ, ν when referring to $\lambda_{\mu,\nu}$ and $\eta_{\mu,\nu}$.

Theorem 4.19. *Let $(S, C, q; >)$ be a college admissions problem. Then $\mu, \nu \in \mathcal{S}(P) \implies \lambda, \eta \in \mathcal{S}(P)$.*

Proof. Assuming that the original college admissions problem is $(S, C, q; >)$, we consider the related marriage market $(S, \widetilde{C}; \widetilde{>})$ and the matchings $\widetilde{\mu}, \widetilde{\nu} \in \mathcal{S}(\widetilde{P})$, corresponding to $\mu, \nu \in \mathcal{S}(P)$, respectively.

By Theorem 3.19, $\widetilde{\lambda} = \widetilde{\mu} \vee_{\widetilde{C}} \widetilde{\nu} \in \mathcal{S}(\widetilde{P})$. Now, by Lemma 4.16, if for any college $c \in C$, $\lambda(c) = \mu(c)$ (the proof for $\lambda(c) = \nu(c)$ is analogous), then $\mu \succeq_c \nu$, and so $\widetilde{\mu}(c^i) \mathrel{\widetilde{\succeq}_{c^i}} \widetilde{\nu}(c^i)$ $\forall c^i$ position of c, and hence $\widetilde{\lambda}(c^i) = \widetilde{\mu}(c^i)$, $\forall c^i$. Therefore, $\forall \sigma \in \mu(c)$, $\exists c^i$ position of c such that $\sigma = \widetilde{\lambda}(c^i)$.

Suppose now that λ is not a matching. That is, $\exists s \in S$, $\exists c_1, c_2 \in C$, $c_1 \neq c_2$, such that $s \in \lambda(c_1) \cap \lambda(c_2)$. By what we have just seen, there must be some positions c_1^i of c_1 and c_2^j of c_2 such that $s = \widetilde{\lambda}(c_1^i) = \widetilde{\lambda}(c_2^j)$, but then $\widetilde{\lambda}$ would not be a matching, and we have a contradiction. Thus, $\lambda \in \mathcal{M}(S, C)$.

Let us see that $\lambda \in \mathcal{S}(P)$. Clearly, it is individually rational, so let us assume it is not in $\mathcal{S}(P)$ by supposing there is a blocking pair $\{c, s\}$. Then, $s >_c \sigma$, for some $\sigma \in \lambda(c)$, and so there must be some position c^i of c such that $\sigma = \widetilde{\lambda}(c^i)$ and $s \mathrel{\widetilde{>}_{c^i}} \widetilde{\lambda}(c^i)$. But now, by Theorem 3.19, $\widetilde{\lambda} \in \mathcal{S}(\widetilde{P}) \implies \widetilde{\lambda}(s) \mathrel{\widetilde{>}_s} c^i \implies \lambda(s) >_s c \implies \{c, s\}$ does not block λ.

To prove $\eta \in \mathcal{S}(P)$, the reasoning is analogous. $\square$

Combining the results given in Theorems 4.17 and 4.19, we have the following corollary, which is the generalisation of Theorem 3.19 for the college admissions problem.

Corollary 4.20. *$\mathcal{S}(P)$ forms a complete and distributive lattice $(L_C, \succeq_C)$ (resp. $(L_S, \succeq_S)$) under the partial order $\succeq_C$ (resp. $\succeq_S$), and the lattice $(L_C, \succeq_C)$ is dual to the lattice $(L_S, \succeq_S)$.*

4.6 Stable matchings and the core

After all the results seen until now, a natural question arises: is the set of stable matchings related in a clear way to the core of $(S, C, q; >)$, like in the marriage problem? We will answer this question after some previous definitions that are very similar to those in the marriage market.

Definition 4.21. Let $(S, C, q; >)$ be a college admissions problem, and let $\mu, \nu \in \mathcal{M}(S, C)$. ν *dominates* μ via a coalition $A \subset S \cup C$ if, $\forall s \in S \cap A$, $\forall c \in C \cap A$,

(a) $c' = \nu(s) \implies c' \in A$,

(b) $s' \in \nu(c) \implies s' \in A$,

(c) $\nu(s) >_s \mu(s)$, and

(d) $\nu(c) >_c \mu(c)$.

Instead, ν *weakly dominates* μ via A if $\forall s \in S \cap A$, $\forall c \in C \cap A$, (a) and (b) are true and,

(c′) $\nu(s) \geq_s \mu(s)$ and $\exists s^* \in A$ such that $\nu(s^*) >_{s^*} \mu(s^*)$, and

(d′) $\nu(c) \geq_c \mu(c)$ and $\exists c^* \in A$ such that $\nu(c^*) >_{c^*} \mu(c^*)$.

The core of the game is defined as in the marriage market, but as we have just defined another relation, the weak domination, we will also define a new core, the core defined by weak domination.

Definition 4.22. The *core* of the college admissions problem $(S, C, q; >)$ is the subset of $\mathcal{M}(S, C)$ formed by all the matchings that are not dominated by any other matching. It is denoted by $\mathcal{C}(P)$.

Definition 4.23. The *core defined by weak domination* of the college admissions problem $(S, C, q; >)$ is the set of matchings that are not weakly dominated by any other matching. It is denoted by $\mathcal{C}_W(P)$.

Trivially, $\mathcal{C}_W(P) \subseteq \mathcal{C}(P)$, because domination implies weak domination, but unlike in the marriage market, in the college admissions problem we have that $\mathcal{C}_W(P) \neq \mathcal{C}(P)$. However, our next proposition states that under our assumptions regarding preferences, the set of stable matchings equals the core defined by weak domination.

Proposition 4.24. *Let $(S, C, q; >)$ be a college admissions problem. If preferences over individuals are strict, then $\mathcal{S}(P) = \mathcal{C}_W(P)$.*

Proof.

$\subseteq$) If $\mu \notin \mathcal{C}_W(P)$, then μ is weakly dominated via a coalition $A \subset S \cup C$ by some other matching $\nu \in \mathcal{M}(S, C)$. Therefore, $\exists a \in A$ such that $\nu(a) >_a \mu(a)$. Now, if μ is individually irrational, then it is obviously unstable, and so $\mu \notin \mathcal{S}(P)$. Suppose then that μ is individually rational, and that there is a $c \in C \cap A$ (the reasoning for $s \in S \cap A$ is even less complicated) such that $\nu(c) >_c \mu(c)$; then, $\exists s \in \nu(c) \backslash \mu(c)$, $\exists \sigma \in \mu(c) \backslash \nu(c)$ such that $s >_c \sigma$. Now, by definition, $s \in \nu(c) \Rightarrow s \in A$, and because $\nu(s) \neq \mu(s)$, the only possibility is that $\nu(s) >_s \mu(s)$. Therefore, μ is blocked via the college-student pair $\{c, s\}$, and $\mu \notin \mathcal{S}(P)$.

$\supseteq$) If $\mu \notin \mathcal{S}(P)$, there are two possible cases: on the one hand, if μ is individually irrational for an agent $x \in S \cup C$, then it is clearly strictly (not simply weakly) dominated by the matching $\nu \in \mathcal{M}(S, C)$ such that $\nu(y) = \mu(y), \forall y \in (S \cup C) \backslash \{x\}$, and $\nu(x) = x$ if $x \in S$, or $\nu(x) = (\mu(x) \backslash \{s\}) \cup \{x\}$ if $x \in C$, where $s \in S$ is unacceptable to x. On the other hand, if μ is unstable via the college-student pair $\{c, s\}$, then $\exists \sigma \in \mu(c)$ such that $s >_c \sigma$; in this case, the matching μ is weakly dominated via the coalition $(\{c\} \cup \{s\} \cup \mu(c)) \backslash \{\sigma\}$ by a matching $\nu \in \mathcal{M}(S, C)$ such that $\nu(s) = c$ and $\nu(c) = (\{s\} \cup \mu(c)) \backslash \{\sigma\}$.

□

Notice that Proposition 4.24 states that there might exist some unstable matching in the core, which is a quite surprising result.

Example 4.25. Let us consider the college admissions problem $(S, C, q; >)$, with $S = \{s_1, s_2, s_3\}$, $C = \{c_1, c_2\}$ and $q = (2, 1)$. Agents are endowed with the following preferences over single agents:

$$\begin{array}{ll} c_1 >_{s_i} s_i >_{s_i} c_2, & \forall i = 1, 2, 3 \\ s_1 >_{c_j} s_2 >_{c_j} s_3 >_{c_j} c_j, & \forall j = 1, 2 \end{array}$$

In this case, there are only the three following individually rational matchings:

$$\mu_1 = \begin{pmatrix} c_1 & c_2 \\ s_1\ s_2 & (c_2) \end{pmatrix}, \mu_2 = \begin{pmatrix} c_1 & c_2 \\ s_1\ s_3 & (c_2) \end{pmatrix}, \text{ and } \mu_3 = \begin{pmatrix} c_1 & c_2 \\ s_2\ s_3 & (c_2) \end{pmatrix}.$$

Moreover, $\mathcal{S}(P) = \mu_1$, since μ_2 is blocked by the pair $\{c_1, s_2\}$ and μ_3 is blocked by the pair $\{c_1, s_1\}$.

However, let us see that $\mu_2 \in \mathcal{C}(P)$. On the one hand, μ_2 is not dominated by μ_3, because there is only one possible coalition, which is $A = \{s_2, s_3, c_1\}$, but $\mu_3(c_1) = \{s_2, s_3\}$

and $\mu_2(s_3) >_{s_3} \mu_3(s_3)$. On the other hand, μ_2 is not strictly dominated by μ_1 because the only possible coalition is $A = \{s_1, s_2, c_1\}$, but $\mu_1(s_1) = c_1 = \mu_2(s_1)$.

Therefore, we have that $\mu_2 \notin \mathcal{S}(P)$, but $\mu_2 \in \mathcal{C}(P)$.

Chapter 5

Alternative mechanisms

As we have mentioned in Chapter 3, the DAA is not the only possible mechanism we can use to find a possible solution for the college admissions problem. There are many problems equivalent to the college admissions problem which are quite common in real life and therefore, the fact that many different mechanisms are used should not come as a surprise. That is why we turn now to review some of them.

Different mechanisms yield different outcomes. For instance, we have seen that the DAA yields a stable matching, the best one for the proposing side of the market. When an administration is faced with a college admissions problem, choosing one mechanism over another depends only on which properties of the outcome are most valued for the particular problem it is trying to solve.

In this chapter, three essentially different mechanisms are presented: we start with the top trading cycles mechanism, then turn to the Boston student assignment and then to serial dictatorship, which forks into two different algorithms depending on some characteristics of each particular problem.

5.1 Top Trading Cycles mechanism

The top trading cycles mechanism (TTC) was developed by David Gale (even though he did not publish it), and first introduced by Shapley and Scarf (1974) [32]. It was build to answer a problem which is slightly different to the college admissions problem: the house allocation problem. An outline of this problem is given before the algorithm. However, the algorithm is adapted in order to find a possible solution to the college admissions problem.

For a clear survey and characterisation of the TTC mechanism, see Dur (2012) [10].

The house allocation problem

The house allocation problem is quite similar to the marriage market; in fact, it is another example of a one-to-one matching problem. It draws an economic model of trading indivisible commodities, such as houses, among agents. There are n agents in the market, and each one of them owns one of these commodities (which we will simply call houses from now on). Agents may trade houses, but none of them can have more than one at the same time. Notice that money has no place in this model, as one item is simply traded for another. Every agent is endowed with a preference relation over the set of houses.

Thus, the elements of this market are:

- $A = \{a_1, \dots, a_n\}$ is the set of agents,
- $H = \{h_1, \dots, h_n\}$ is the set of houses, and
- $\forall a \in A$, there is a strict preference relation over houses $>_a$, namely a total order over the set H.

Then, to state that agent $a \in A$ prefers house h to house h', we write that $h >_a h'$. The set of all preference relations is $> := (>_a)_{a \in A}$, which is called a preference profile. Thus, the market is denoted by the triple $(A, H; >)$.[1]

The aim of the market is to redistribute the n commodities taking into account these preferences. It is easy to see that the main difference between this model and the marriage model is that one side of the market (houses) has no say in the matter, and thus its elements are not strategic players. Another remarkable difference is that $|A| = |H|$, always.

The outcome of the market after the necessary trades are made is called an *allocation*.

Definition 5.1. An *allocation* of the house allocation problem $(A, H; >)$ is a bijective map $\alpha : A \to H$.

Just as the marriage or the college admissions problem, the house allocation problem $(A, H; >)$ can be seen as a game $(\mathcal{I}, S, u)$, with $\mathcal{I} = A$, $S = \mathcal{P}(A)$, $\mathcal{P}(A)$ being the set of all possible preference profiles, and $u = u_1 \times \ldots \times u_n$, with u_i any utility function on the set of all possible allocations $\forall i = 1, \dots, n$, which is compatible with the preference relation $>_i$.

Now that we have explained the problem for which the Top Trading Cycles mechanism was first introduced, we will present the algorithm. The following TTC mechanism is an adaptation from Abdulkadiroğlu and Sönmez (2003) [2] for our particular problem.[2] Let $(S, C, q; >)$ be a college admissions problem, with the set of students $S = \{s_1, \dots, s_m\}$ and the set of colleges $C = \{c_1, \dots, c_n\}$, with quotas $q = (q_1, \dots, q_n)$. Students and colleges are endowed with the strict preference profile $P = P(s_1) \times \ldots \times P(s_m) \times P(c_1) \times \ldots \times P(c_n)$.

Step 0. We define, for $j = 1, \dots, n$, counters $i_j := q_j$. We also consider the graph Γ_0 with nodes $S \cup C$ and no edges.

Step 1. (a) Every node $s \in S$ points to her most preferred college $c \in C$ and every node $c \in C$ points to its most preferred student $s \in S$. That is, we consider directed edges from one set to the other and vice-versa.

[1] Actually, it could simply be denoted by the pair $(A; >)$, as H depends only on A.

[2] In fact, in [2], the problem considered is the *school choice problem*, which is midway between house allocation and college admissions, because schools have several seats to offer, but their preferences are more like priorities of students, and so schools are not strategic agents.

(b) If there is some cycle $(c_0, s_0, \dots, c_k, s_k, c_0)$,[3] student s_j is assigned college c_j, and the counter i_j is reduced by one, for $j = 0, \dots, k$. We consider the new graph with no edges Γ_1, with nodes $S_1 = S \backslash \{s_0, \dots, s_k\}$ and $C_1 = \{c_j \in C; i_j \neq 0\}$.

$\vdots$

Step k. (a) Consider Γ_{k-1}. Every node $s \in S_{k-1}$ points to its most preferred college $c \in C_{k-1}$ and every node $c \in C_{k-1}$ points to its most preferred student $s \in S_{k-1}$.

(b) If there is some cycle $(c_0, s_0, \dots, c_l, s_l, c_0)$,[4] student s_j is assigned college c_j, and the counter i_j is reduced by one, for $j = 0, \dots, l$. If $S_k, C_k = \varnothing$, stop. Otherwise, consider the new graph with no edges Γ_k, with nodes $S_k = S_{k-1} \backslash \{s_0, \dots, s_l\}$ and $C_k = \{c_j \in C_{k-1}; i_j \neq 0\}$.

The outcome of this TTC mechanism is an element $\mu \in \mathcal{M}(S, C)$. However, this mechanism is not equivalent to de DAA, because μ might not be in $\mathcal{S}(P)$ (there could be blocking pairs).

Nevertheless, μ is strategy-proof and Pareto-efficient; that is, truth-telling is a dominant strategy for all agents (unlike in the DAA), and no other matching gives a more preferred assignment to at least one agent without leaving someone else worse than they were at μ.

In view of this, it is clear that when faced with the decision to choose between the DAA and the TTC mechanism, any administration should study what properties of the outcome are preferred: both mechanisms can be strategy-proof for students,[5] and so the administration does not need to worry about families lying or playing strategically, but if Pareto-optimality is more important than stability, the TTC should be chosen; otherwise, the DAA should be chosen.

We would also like to highlight that an extension of the TTC mechanism, the Top Trading Cycles and Chains algorithm, is currently used in kidney exchange. We have no space to explain this case, but for the interested reader we recommend Roth and Sönmez (2004) [31].

5.2 Boston Student Assignment

The Boston Student Assignment mechanism is not quite as interesting, mathematically speaking, as the two previous algorithms, but it is the one that has been used to assign public schools in Catalonia and public nursery schools in Barcelona until this year, and so it is important for us to consider it as well.[6] Obviously, it was the one used in Boston public schools too. In Abdulkadiroğlu et al. (2005) [1], the situation of the public school system in Boston in 2005 is concisely explained.

In this algorithm, there is a slightly different approach to agents' preferences. Clearly, if we reformulate the house allocation problem in order to obtain more than one room in every house (that is, each house has now a quota), we obtain a new problem, often called

[3] If there is more than one cycle, we consider only the longest one.

[4] Just as in Step 1, if there is more than one cycle, only the longest one is considered.

[5] Recall that the DAA is strategy-proof for all agents in S when students are the ones proposing.

[6] The new algorithm is a variation of the Serial Dictatorship, which we analyse in the next section.

the *student choice problem*. In this problem, there are agents –students– and schools, and each school has a limited number of seats to offer. But schools are not agents: they cannot choose their preferences, because every student has a priority number for each school. These priority numbers are usually determined by a central administration and are based on characteristics such as vicinity to the school or having a sibling studying there already, and thus schools do not have the opportunity to misrepresent them and play strategically.

That being said, there is a clear isomorphism between school choice and college admissions, and one can treat these school priorities as if they were indeed school preferences. Therefore, that is exactly what will be done hereafter.

Given the college admissions problem $(S, C, q; >)$, with $S = \{s_1, \dots, s_m\}$, $C = \{c_1, \dots, c_n\}$ and preference profile $P = P(s_1) \times \dots \times P(s_m) \times P(c_1) \times \dots \times P(c_n)$, where $P(s_i)$ is the preference ranking of colleges student s_i submits, such that $|P(s_i)| \leq n$, $\forall s_i \in S$, and $P(c_j)$ is the priority order of college c_j over S, $\forall c_j \in C$.

Now, we will adapt the Boston mechanism to the problem $(S, C, q; >)$. Observe that because we consider it from the very beginning, the first thing this algorithm does is to determine the preference profile P.

Step 0. (a) For every $c_i \in C$, a priority ordering $P(c_i)$ over the elements of S is determined, based on administrative policies.

(b) Every student $s \in S$ submits a preference list $P(s)$ stating her preferences over C.

Step 1. The first choices of the students are considered: every college $c_i \in C$ considers all the students who have ranked it as their first choice, and taking into account its priority ordering over $S \cup \{c_i\}$, it accepts these students one by one until either its quota q_i is filled or there are no more students who have chosen it as their top choice.

$\vdots$

Step k. In this step, only the students who have not yet been assigned to any college are considered, as they have been rejected by every college ranked above the k^{th} place. Seats are assigned likewise in Step 1, but considering only students' k^{th} choices.

The Boston mechanism stops when either every student is assigned to a school or every student who is not yet assigned to any school has no more schools on her list. If this is the case, the administration assigns these students to a random school with unfilled quota.

Despite it being widely used, the Boston mechanism has some severe weaknesses which may lead to serious implications for the central administration taking care of the assignment:

- It does not yield a stable outcome, and so many families could become quite unhappy, even to the point of bringing suits against the administration.
- It does not yield a Pareto-optimal outcome, and so it is not an efficient algorithm.
- Clearly, the previous two points imply that families have legitimate reasons to misrepresent their stated preferences, that is, to play strategically, because they can easily benefit from it. Usually, these misrepresentations consist of ranking first in

their list the school for which they have the highest priority, even if it is not their favourite one. This should also be a problem for the central administration, because the lists submitted by students, if true, provide invaluable information that could be used to make decisions.

In [2], there is a brief explanation of the three different mechanisms we have seen so far, as well as a clear comparative with the most important features each one of them has and does not have.

5.3 Serial Dictatorship

The last mechanism we will describe is the serial dictatorship mechanism. This mechanism, just as the top trading cycles mechanism, was first aimed to solve the house allocation problem. The key aspect of this mechanism is that, having specified an order of agents on one side of the market, the first agent on that side chooses her top choice, the second one her top choice among those that are still available, and so on until the last agent remaining.

This mechanism has been used to assign colleges in Turkey or China, for example;[7] in these countries, the process of college admissions is centralised. This means that, just as in the school choice problem, colleges are not strategic agents.

This specific problem is often called *student placement problem*. It differs from the school choice problem in an important aspect: whilst in the school choice problem priorities for each school are determined by an exogenous administration, in the student placement problem students are ranked based on an exam score.

The elements of the student placement problem are the following:

- a set of students $S = \{s_1, \dots, s_m\}$,
- a set of colleges $C = \{c_1, \dots, c_n\}$,
- a quota vector $q = (q_1, \dots, q_n)$, with $q_i \in \mathbb{N}$, $\forall i \in \{1, \dots, n\}$,
- a preference profile for students $>_S = (>_s)_{s \in S}$, where $\forall s \in S$, $>_s$ is a strict preference relation of student s over $C \cup \{s\}$.
- a finite set of *categories* for colleges K,
- an *exam score profile* for students, $e = (e^s)_{s \in S}$, such that for every $s \in S$, $e^s = (e^s_k)_{k \in K}$ with $e^s_k \in (0, +\infty)$ being the exam score of student s in the category $k \in K$,[8] and
- a *type function* that maps each college to a category type, $t : C \to K$.

We can see that there are more elements in this problem than in any of the others we have seen. However, we cannot assert immediately that college $c \in C$ has a preference

[7] See [4] and [35].

[8] There is a rigid limitation to this problem: we suppose that $\nexists s' \in S \backslash \{s\}$ such that $e^s_k = e^{s'}_k$. However, whenever this problem arises, ties can be broken in different ways such as the students' ages or the alphabetical order of their names.

relation over $S \cup \{c\}$, and therefore this is obviously not the typical college admissions problem. The motivation to introduce categories and therefore exam score profiles and a type function is that, even though all colleges consider a unique exam which is taken by all students, each college may give more importance to one part of the exam or another.[9]

Formally, this means that each college $c \in C$ admits students according to their exam score in category $t(c)$. As there is a unique category for each college, we can consider a preference relation for every c, $>_c$, based solely on the ranking in its category $t(c)$; now, we can denote a student placement problem by $(S, C, q, >, e, t)$, with $> := (>_x)_{x \in S \cup C}$.

When $|K| = 1$, we speak of *one skill category serial dictatorship*,[10] and when $|K| > 1$, we speak of *multi-category serial dictatorship*. We start by displaying the one skill category mechanism and then turn to the multi-category mechanism.

One Skill Serial Dictatorship Mechanism

Let us consider the student placement problem $(S, C, q, >, e, t)$. The algorithm is the following:

Step 0. We define, for $j = 1, \ldots, |C|$, counters $i_j := q_j$.

Step 1. The highest priority agent, that is, $s_1 := \arg\max_{s \in S}(e^s)$, is assigned to her preferred choice under $>_{s_1}$. We define $S_1 := S \backslash \{s_1\}$, and the counter corresponding to the college to which s_1 has been assigned is lowered by one.

$\vdots$

Step k. The k^{th} highest priority agent, $s_k := \arg\max_{s \in S_{k-1}}(e^s)$, is assigned to her top choice under $>_{s_k}$ among those that are still available, that is, the ones with $i_j \neq 0$. If no such college exists, s_k remains self-matched. We define $S_k := S_{k-1} \backslash \{s_k\}$ and the counter corresponding to the college to which s_k has been assigned is lowered by one. If $S_k = \varnothing$ or $i_j = 0 \ \forall j$, the algorithm stops.

Multi-Category Serial Dictatorship Mechanism

Let us introduce the *Multiple-Category Serial Dictatorship* algorithm for the student placement problem $(S, C, q, >, e, t)$.

Step 1. (a) For every category $k \in K$, we consider the ranking induced by the exam scores $(e^s_k)_{s \in S}$, and implement the induced one skill serial dictatorship. This leads to a *tentative student placement*.[11]

(b) $\forall s \in S$, we construct a new preference relation $>^1_s$ as follows: if s is not assigned to more than one college, $>^1_s := >_s$; otherwise, $>^1_s$ is the preference relation obtained from $>_s$ by moving the option of remaining self-matched, s,

[9]When we consider the university entrance procedure in Spain, which is based mainly on the *examen de selectivitat*, it is reasonable that different faculties ponder differently, for instance, the students' English score.

[10]In this case, all colleges have the exact same preferences over S.

[11]Notice that this is not necessarily a matching, because a student could be assigned to more than one college.

right after s's favourite college among those that are tentatively assigned to her, and keeping everything else unchanged. We obtain a new preference profile (restricted to S) $>^1 := (>^1_s)_{s \in S}$.

$\vdots$

Step k. Proceed as in Step 1, and construct $>^k$ from $>^{k-1}$. If no student is assigned more than one college, the algorithm stops.

Our next theorem shows the relationship between this last algorithm and the DAA with colleges proposing. Its proof can be found in [4].

Theorem 5.2. (Balinski and Sönmez, 1999) *Let us consider the student placement problem* $(S, C, q, >, e, t)$*. Then, the multi-category serial dictatorship algorithm is equivalent to the deferred acceptance algorithm with colleges proposing for the related college admissions problem* $(S, C, q; >)$*.*

Notice that this implies that multi-category serial dictatorship, just as deferred acceptance, is not Pareto-efficient, but yields the stable matching μ_C.

Final Remarks

As we have seen throughout this work, examples of two-sided matching markets are quite common in our daily lives, yet almost unknown to the public.

The main purpose of the current study was to set the foundations of matching theory in the most mathematically correct possible manner, and I think this goal has been fulfilled to a great extent. In order to do so, we may have explained the marriage problem more extensively than we thought we would; indeed, the first title we had thought of was "The College Admissions Problem", but it was clear, after having read most of the bibliography, that the marriage problem had to have a privileged place in the final draft, as it is the basis from which other two-sided matching markets are built.

In this study, we have focused on the marriage problem and the college admissions problem, and have only quickly mentioned the house allocation problem, the student choice problem, the student placement problem and kidney exchange. However, there are many more matching markets, such as the labour market, or even dating apps.

We have seen how agents in these markets should act interdependently, depending on the properties of the outcome resulting from the algorithm used by the central administration that supervises the market; properties such as stability, strategy-proofness or Pareto-dominance are studied.

While doing this work, we have tried to investigate how school choice functions in Catalonia, and have completely failed to find straight to the point information; it is hard to understand why the mechanism used is not made public, and I hope that the Generalitat becomes more transparent, at least in this aspect, in the near future.

To conclude, it has been a pleasant surprise to discover a field of study that, by means of mathematical arguments, explains such natural and common situations of the modern world.

Bibliography

[1] A. Abdulkadiroğlu, P.A. Pathak, A.E. Roth, T. Sönmez, *The Boston Public School Match*, The American Economic Review, **95**, no. 2, (2005), 368–371.

[2] A. Abdulkadiroğlu, T. Sönmez, *School Choice: A Mechanism Design Approach*, The American Economic Review, **93**, no. 3, (2003), 729–747.

[3] M. Baïou, M. Balinski, *Erratum: The Stable Allocation (or Ordinal Transportation) Problem*, Mathematics of Operations Research, **27**, no. 4, (2002), 662–680.

[4] M. Balinski, T. Sönmez, *A Tale of Two Mechanisms: Student Placement*, Journal of Economic Theory, **84**, no. 1, (1999), 73–94.

[5] C. Blair, *The Lattice Structure of the Set of Stable Matchings with Multiple Partners*, Mathematics of Operations Research, **13**, no. 4, (1988), 619–628.

[6] Y. Chen, T. Sönmez, *School Choice: an Experimental Study*, Journal of Economic Theory, **127**, no. 1, (2006), 202–231.

[7] V.P. Crawford, E.M. Knoer, *Job Matching with Heterogeneous Firms and Workers*, Econometrica, **49**, (1981), 437-–450.

[8] E. Drgas-Burchardt, Z. Świtalski, *A Number of Stable Matchings in Models of the Gale–Shapley Type*, Discrete Applied Mathematics, **161**, no. 18, (2013), 2932–2936.

[9] L.E. Dubins, D.A. Freedman, *Machiavelli and the Gale-Shapley Algorithm*, The American Mathematical Monthly, **88**, no. 7, (1981), 485–494.

[10] U.M. Dur, *A Characterization of the Top Trading Cycles Mechanism for the Shcool Choice Problem*, MRPA Paper, (2012).

[11] D. Fudenberg, J. Tirole, *Game Theory*, MIT Press, 1991.

[12] D. Gale, *The Theory of Linear Economic Models*, Chicago: The University of Chicago Press, 1960.

[13] D. Gale, *The Two-Sided Matching Problem: Origin, Development and Current Issues*, International Game Theory Review, **3**, no. 2, (2001), 237–252.

[14] D. Gale, L.S. Shapley, *College Admissions and the Stability of Marriage*, The American Mathematical Monthly, **69**, no. 1, (1962), 9–15.

[15] D. Gale, M. Sotomayor, *Some Remarks on the Stable Matching Problem*, Discrete Applied Mathematics, **11**, no. 11, (1983), 223–232.

[16] R. Gardner, *Juegos para empresarios y economistas*, Barcelona: Antoni Bosch, 1996.

[17] J. González-Díaz, I. García-Jurado, M.G. Fiestras-Janeiro, *An Introductory Course on Mathematical Game Theory*, AMS Graduate Studies in Mathematics, Vol. 115, 2010.

[18] D.E. Knuth, *Mariages Stables*, Montréal: Les Presses de l'Université de Montréal, 1976.

[19] F. Kojima, P.A. Pathak, A.E. Roth, *Matching With Couples: Stability and Incentives in Large Markets*, The Quarterly Journal of Economics, **128**, no. 4, (2013), 1585–1632.

[20] J.F. Nash, *Equilibrium points in n-person games*, Proceedings of the National Academy of Sciences, **36**, (1950), 48–49.

[21] "The Prize in Economic Sciences 2012", `http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2012`, accessed: 2018-06-01.

[22] J. Pérez, J.L. Jimeno, E. Cerdá, *Teoría de Juegos*, Madrid: Garceta, 2013.

[23] A.E. Roth, *The Economics of Matching: Stability and Incentives*, Mathematics of Operations Research, **7**, no. 4, (1982), 617–628.

[24] A.E. Roth, *The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory*, Journal of Political Economy, **92**, no. 6, (1984), 991–1016.

[25] A.E. Roth, *Misrepresentation and Stability in the Marriage Problem*, Journal of Economic Theory, **34**, no. 2, (1984), 383–387.

[26] A.E. Roth, *The College Admissions Problem is Not Equivalent to the Marriage Problem*, Journal of Economic Theory, **36**, no. 2, (1985), 277–288.

[27] A.E. Roth, *Deferred Acceptance Algorithms: History, Theory, Practice and Open Questions*, International Journal of Game Theory, **36**, no. 3, (2008), 537–569.

[28] A.E. Roth, M. Sotomayor, *The College Admissions Problem Revisited*, Econometrica, **57**, no. 3, (1989), 559–580.

[29] A.E. Roth, M. Sotomayor, *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis*, Cambridge University Press, 1990.

[30] A.E. Roth, *The economist as engineer: Game Theory, Experimentation, and Computation as tools for design economics*, Econometrica, **70**, no. 4, (2002), 1341–1378.

[31] A.E. Roth, T. Sönmez, M.U. Ünver, *Kidney Exchange*, The Quarterly Journal of Economics, **119**, no. 2, (2004), 457–488.

[32] L. Shapley, H. Scarf, *On Cores and Indivisibility*, Journal of Mathematical Economics, **1**, no. 1, (1974), 23–37.

[33] L.S. Shapley, M. Shubik, *The Assignment Game I: The Core*, International Journal of Game Theory, **1**, (1972). 111-–130.

[34] T. Sönmez, M.U. Ünver, *Matching, Allocation, and Exchange of Discrete Resources*, J. Benhabib, A. Bisin, and M. Jackson (eds.), Handbook of Social Economics, Vol. 1A, North-Holland (2011), 781–852.

[35] M. Zhu, *College Admissions in China: A Mechanism Design Perspective*, **30**, (2014), 618–631.