

Some things couples always wanted to know about stable matchings (but were afraid to ask)*

Bettina Klaus[†] Flip Klijn[‡] Jordi Massó[§]

November 2005

Abstract: In this note we study the National Resident Matching Program (NRMP) algorithm in the US market for physicians. We report on two problems that concern the presence of couples, a feature explicitly incorporated in the new NRMP algorithm (*cf.* Roth and Peranson, 1999). First, we show that the new NRMP algorithm may not find an existing stable matching, even when couples' preferences are 'responsive,' *i.e.*, when Gale and Shapley's (1962) deferred acceptance algorithm (on which the old NRMP algorithm is based) is applicable. Second, we demonstrate that the new NRMP algorithm may also be manipulated by couples acting as singles.

Keywords: Matching; Stability; Couples

JEL classification: C78; D78; J41

1 Introduction

In many countries, the proportion of women attending college has steadily been increasing during the last decades. Thus, it is not surprising that the number of couples searching jointly for a job in the same labor market has been increasing as well. In this paper we deal with a specific US labor market that, because of its history and development, is a bench mark for several other matching markets in the US and Canada. Each year thousands of medical school graduates seek their first employment through a centralized matching process, the National Resident Matching Program (NRMP).¹ This clearinghouse was initiated in the 1950's in response to persistent failures to organize the market in a timely and orderly way by decentralized means. Roth (1984) would

*We thank Howard Petith, Alvin Roth, William Thomson, and an anonymous referee for helpful comments and suggestions. B. Klaus's and F. Klijn's research was supported by Ramón y Cajal contracts of the Spanish *Ministerio de Ciencia y Tecnología*. The main part of F. Klijn's work was supported by a Marie Curie Fellowship of the European Community programme "Improving Human Research Potential and the Socio-economic Knowledge Base" under contract number HPMF-CT-2001-01232, carried out at the Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona. J. Massó's research was partially supported by Research Grant 2001SGR-00162 from the *Department d'Universitats, Recerca i Societat de la Informació (Generalitat de Catalunya)*. The work of the authors was partially supported by Research Grant BEC2002-02130 from the Spanish *Ministerio de Ciencia y Tecnología* and by the Barcelona Economics Program of CREA.

[†]Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands, e-mail: B.Klaus@algec.unimaas.nl

[‡]Corresponding author. Institut d'Anàlisi Econòmica (CSIC), Campus UAB, 08193 Bellaterra (Barcelona), Spain; Tel. (34) 93 580 6612; Fax. (34) 93 580 1452; e-mail: fkljnj@pareto.uab.es

[§]CODE and Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain; e-mail: jordi.massou@uab.es

¹See Roth (1984), Roth and Sotomayor (1990), and Roth and Peranson (1999).

later explain its success by showing that the clearinghouse in fact employed the hospital-optimal variant of the so-called deferred acceptance algorithm, developed by Gale and Shapley (1962) who at the time were not aware of the relation with the NRMP.

Around the mid 1970's voluntary and orderly participation in the NRMP started to drop. What happened then was that a growing number of couples in need of two positions in the same vicinity left the centralized market and started to negotiate directly with hospitals (see Checker, 1973). As a consequence, the labor market became, just as before the 1950's, prone to chaos and dissatisfaction on all sides. A hypothesis offered by Roth (1984) is that the chaotic conditions reflect the instability of the matching procedure. If couples and hospitals find it profitable to make their own arrangements outside of the matching program it must be that the matching procedure is unstable with respect to couples. This indeed turned out to be the case. In the mid 1990's a crisis of confidence² in the matching procedure on the applicants side finally induced the NRMP Board of Directors to design a new algorithm. Apart from recovering students' confidence by favoring their side, the algorithm was also meant to deal with couples in an appropriate manner.³ The first match with the new algorithm was carried out in 1998. Roth and Peranson (1999) describe how the new algorithm was designed. Furthermore, using computational simulations and analyzing previous data, they show that the new algorithm is to be expected to perform well in practice. Roth (2002) gives a more recent review of the redesign of the NRMP algorithm in the context of analyzing the 'engineering aspects' of economic design. A nice overview of how the new algorithm was designed to address the problems that occur in the presence of couples is given in Roth (2002, Section 2.4.1).

In this note, we address two issues that are of importance for couples that participate in the new NRMP algorithm. The first issue concerns the stability of the matching obtained by the NRMP algorithm. A matching is 'stable' if (a) each couple and each hospital have acceptable matches, and (b) no couple and no pair of hospitals prefer a mutual match to the ones that have been assigned. Roth (1984) and Sotomayor (unpublished note) independently demonstrate the possibility of instability in the presence of couples. However, Klaus and Klijn (2005a) show that for a natural preference domain for couples, namely the domain of 'responsive' preferences, there is always a stable matching that moreover can be found by applying Gale and Shapley's (1962) deferred acceptance algorithm. In Example 3.1 we construct a matching market in which couples have responsive preferences and where there is a unique stable matching, which hence can be easily obtained by Gale and Shapley's (1962) deferred acceptance algorithm. Surprisingly, we still find that the new NRMP algorithm may cycle, that is, not find the stable matching.

Next, couples may wonder if they should apply to the NRMP as a couple or as two seemingly independent applicants. With our second couples market (Example 3.2) we show that the new NRMP algorithm may indeed be prone to strategic manipulation by the members of a couple pretending to be singles, *i.e.*, a couple may be better off by applying as two separate applicants.

Both our couples markets are realistic in the sense that (a) couples' preferences can be easily explained in terms of individual preferences and distances between hospitals, and (b) hospitals have the same preferences over students (which can be justified by final grades or test scores).

The paper is organized as follows. In Section 2, we introduce our matching model with couples. In Section 3, we present the two potential problems of the new NRMP algorithm. Finally, in Section 4 we discuss the difficulties to avoid these problems.

²Many students believed that the matching was not conducted in their best interest and that possibilities for strategic manipulations existed; see Roth and Peranson (1999).

³In fact, three other issues that are not relevant in the present discussion were addressed as well. See Roth (2002, p. 1355) for details.

2 Matching with couples: the model

We describe a model with 4 hospitals and 2 pairs of students; $H = \{h_1, h_2, h_3, h_4\}$, $S = \{s_1, s_2, s_3, s_4\}$, and $C = \{c_1, c_2\} = \{(s_1, s_2), (s_3, s_4)\}$ are the sets of hospitals, students, and couples, respectively. Each hospital has exactly one position to be filled. Our definitions and results can easily be adapted to more general situations that include other couples as well as single agents and hospitals with multiple positions.

Hospitals' preferences: Each hospital $h \in H$ has a strict, transitive, and complete preference relation \succeq_h over the set of students S and the prospect of having its position unfilled, denoted by \emptyset ; for instance, $P(h) = s_4, s_2, \emptyset, s_1, s_3$ indicates that hospital h prefers student s_4 to s_2 , and considers students s_1 and s_3 to be unacceptable. Let $P^H = \{P(h)\}_{h \in H}$.

Students' preferences: Similarly, each student $s \in S$ has an individual strict, transitive, and complete preference relation \succeq_s over the set of hospitals and the prospect of being unemployed, denoted by u . We assume that these individual preferences are the preferences a student has if he is single. For instance, $P(s) = h_1, h_2, h_3, h_4, u$ indicates that student s prefers h_i to h_{i+1} for $i = 1, 2, 3$ and prefers being employed to being unemployed. Let $P^S = \{P(s)\}_{s \in S}$.

Couples' preferences: Each couple $c \in C$ has a strict, transitive, and complete preference relation \succeq_c over all possible combination of ordered pairs of (different) hospitals and the prospect of being unemployed. We denote a generic ordered pair of hospitals by (h_p, h_q) , where h_p and h_q indicate either a hospital or being unemployed. For instance, $P(c) = (h_4, h_2), (h_3, h_4), (h_4, u)$, etc., indicates that couple $c = (s_1, s_2)$ prefers s_1 and s_2 being matched to h_4 and h_2 , respectively, to being matched to h_3 and h_4 , respectively, and so on. Let $P^C = \{P(c)\}_{c \in C}$.

Singles and couples markets: Now, the standard one-to-one two-sided matching market with single students, or *singles market* for short, is denoted by (P^H, P^S) . Since singles markets and some of the classical results for singles markets are well-known, for a detailed description we refer to Roth and Sotomayor (1990). We define a one-to-one matching market with couples, or a *couples market* for short, by (P^H, P^C) .

Matchings: A *matching* μ for a couples market (P^H, P^C) is an assignment of students and hospitals such that each student is assigned to at most one hospital in H or to u (which can be assigned to multiple students), each hospital in H is assigned to at most one student or to \emptyset (which can be assigned to multiple hospitals), and a student is assigned to a hospital if and only if the hospital is assigned to the student. A matching μ is completely described by the list $\mu(H) = \mu(h_1), \mu(h_2), \mu(h_3), \mu(h_4)$ of students in S or \emptyset matched to hospitals h_1, h_2, h_3, h_4 .

Stability: Finally, we define stability for couples markets (see Roth and Sotomayor, 1990). First, for a matching to be stable, it should always be better for students (one or both members in a couple) to accept the position(s) offered by the matching instead of voluntarily choosing unemployment and for hospitals it should always be better to accept the student assigned by the matching instead of leaving the position unfilled. A matching μ is *individually rational* if

(i1) for all $c = (s_k, s_l)$, $(\mu(s_k), \mu(s_l)) \succeq_c (\mu(s_k), u)$, $(\mu(s_k), \mu(s_l)) \succeq_c (u, \mu(s_l))$,
and $(\mu(s_k), \mu(s_l)) \succeq_c (u, u)$;

(i2) for all $h \in H$, $\mu(h) \succeq_h \emptyset$.

Second, if one partner in a couple can improve the given matching for the couple by switching to another hospital such that this hospital is better off as well, then we would expect this mutually beneficial trade to be carried out, rendering the given matching unstable. A similar statement

holds if both students in the couple can improve. For a given matching μ , $(c = (s_k, s_l), (h_p, h_q))$ is a *blocking coalition* if

(b1) $(h_p, h_q) \succ_c (\mu(s_k), \mu(s_l))$;

(b2) $[h_p \in H \text{ implies } s_k \succeq_{h_p} \mu(h_p)]$ and $[h_q \in H \text{ implies } s_l \succeq_{h_q} \mu(h_q)]$.

A matching is *stable* if it is individually rational and if there are no blocking coalitions.⁴

Roth (1984) and Sotomayor (unpublished note) demonstrate that stable matchings may not exist in the presence of couples. Klaus and Klijn (2005a,b) prove existence for couples market where couples' preferences are 'responsive' by applying Gale and Shapley's (1962) deferred acceptance algorithm to the 'associated individual preferences.' A couple's preferences are responsive if the unilateral improvement of one partner's job is considered beneficial for the couple as well. For instance, responsive preferences may reflect situations where couples search for jobs in the same metropolitan area (if one partner switches to a job he/she prefers and the couple can still live together, then the couple is better off).⁵

Responsive preferences: Couple $c = (s_k, s_l)$ has *responsive* preferences if there exist *associated individual preferences* \succeq_{s_k} and \succeq_{s_l} such that for all $h_p, h_q, h_r \in H \cup \{u\}$, $[h_p \succ_{s_k} h_r \text{ implies } (h_p, h_q) \succ_c (h_r, h_q)]$ and $[h_p \succ_{s_l} h_r \text{ implies } (h_q, h_p) \succ_c (h_q, h_r)]$. If these preferences \succeq_{s_k} and \succeq_{s_l} exist, then they are unique.

3 Couples and the new NRMP algorithm

The formal and complete description of the new algorithm used by the NRMP is outside the scope of this note. Essentially, the new algorithm adapts (from the original one-to-one model without couples) Roth and Vande Vate's (1990) dynamic process to find stable matchings. Flowchart 1 in the Appendix describes concisely the parts of the new Applicant Proposing Couples Algorithm (APCA) used by the NRMP that are relevant for applying the algorithm to Examples 3.1 and 3.2 below.⁶

First, we demonstrate that even for responsive preferences the algorithm might cycle without selecting a stable matching. This example is particularly interesting because the unique stable matching could be easily found using the deferred acceptance algorithm.

Example 3.1 *The New NRMP Algorithm may cycle for Responsive Preferences*
 Consider the couples market (P^H, P^C) given in Table 1. Students' individual preferences P^S equal $P(s_1) = P(s_2) = h_1, h_2, h_3, h_4, u$, $P(s_3) = h_2, h_1, h_3, h_4, u$, and $P(s_4) = h_3, h_4, h_2, h_1, u$. Note that the couples' preferences are responsive: any unilateral improvement of one partner's job is considered beneficial for the couple as well. Moreover, hospitals have identical preferences over students, which can be easily justified if hospitals rank students according to final grades or other test scores. The unique stable matching for (P^H, P^C) is $\mu(H) = s_2, s_3, s_1, s_4$ (Appendix, Table 2), which, because of responsiveness, is the outcome of the deferred acceptance algorithm.

⁴In order to keep notation as simple as possible, we allow some redundancy in the definition of stability with respect to (i1) and (b1).

⁵For a discussion of results related to responsiveness in couples markets we refer to Klaus and Klijn (2005a,b).

⁶The complete flowchart describing the APCA algorithm that we used to derive Flowchart 1 was kindly made available to us by Alvin Roth.

Next, we apply the Applicant Proposing Couples Algorithm to this couples market.⁷ Suppose that couple (s_3, s_4) is at the top of the stack (a symmetric process occurs if instead couple (s_1, s_2) is at the top of the stack). The algorithm starts with the empty matching $\mu^0(H) = \emptyset, \emptyset, \emptyset, \emptyset$ and cycles over the unstable matchings $\mu^I(H) = \emptyset, s_3, s_4, \emptyset$, $\mu^{II}(H) = s_1, s_2, \emptyset, \emptyset$, $\mu^{III}(H) = s_3, \emptyset, s_4, \emptyset$, $\mu^{IV}(H) = s_2, s_1, \emptyset, \emptyset$, and finally back to $\mu^I(H)$. \diamond

P^H				P^C	
h_1	h_2	h_3	h_4	s_1s_2	s_3s_4
s_2	s_2	s_2	s_2	h_1h_2	h_2h_3
s_3	s_3	s_3	s_3	h_1h_3	h_2h_4
s_1	s_1	s_1	s_1	h_1h_4	h_2h_1
s_4	s_4	s_4	s_4	h_2h_1	h_1h_3
\emptyset	\emptyset	\emptyset	\emptyset	h_2h_3	h_1h_4
				h_2h_4	h_1h_2
				h_3h_1	h_3h_4
				h_3h_2	h_3h_2
				h_3h_4	h_3h_1
				h_4h_1	h_4h_3
				h_4h_2	h_4h_2
				h_4h_3	h_4h_1
				\vdots	\vdots

Table 1: Responsive preferences for which the new NRMP algorithm cycles

Finally, we illustrate the possibility that, if the APCA is used, a couple may obtain a better pair of positions by registering as single students rather than as a couple.

Example 3.2 Pretending to be Single may be Beneficial

Consider the couples market (P^H, P^C) where hospitals h_1 and h_2 are located in one city and hospitals h_3 and h_4 in some other, distant city. Students and couples have the same preferences over hospitals: $P(s) = h_1, h_2, h_3, h_4, u$ for each $s \in S$ and $P(s_1, s_2) = P(s_3, s_4) = (h_1, h_2), (h_2, h_1), (h_3, h_4), (h_4, h_3), \dots$ (the tail can be anything). In other words, a couple would look for two positions in another market before accepting two positions located in different cities. The hospitals' preferences over students are $P(h) = s_1, s_3, s_4, s_2, \emptyset$ for every $h \in H$.

Assume first that the four students register as couples, and couple (s_1, s_2) is at the top of the stack. Then, the APCA yields the matching $\tilde{\mu}(H) = s_1, s_2, s_3, s_4$.

However, if s_3 and s_4 register as single students and, as a consequence, the order in the stack changes to $s_3, s_4, (s_1, s_2)$, then the algorithm yields the matching $\hat{\mu}(H) = s_3, s_4, s_1, s_2$. At this matching couple (s_3, s_4) is strictly better off than at matching $\tilde{\mu}$.⁸ \diamond

⁷For a step by step description of the execution of the algorithm for this example as well as Example 3.2 see <http://pareto.uab.es/fklijn/papers/couplesnrm.htm> or contact any of the authors.

⁸If instead the order in the stack changes to $s_4, s_3, (s_1, s_2)$, then the algorithm produces the matching $\bar{\mu}(H) = s_4, s_3, s_1, s_2$, in which couple (s_3, s_4) is also strictly better off than at matching $\tilde{\mu}$.

4 Two final remarks

One may wonder whether the two potential problems we describe in the previous section can be fixed. First, consider the problem of finding a stable matching whenever one exists. If couples' preferences are responsive, a stable matching exists and can be found by applying the DA-algorithm to the associated individual preferences. (Note that the DA-algorithm is polynomially bounded in time.) However, if couples' preferences are not responsive, then we cannot construct associated individual preferences.⁹ Furthermore, in any attempt to construct associated individual preferences for a non-responsive preference relation, one might neglect an important complementarity, *e.g.*, the wish to live together in the same city, that is crucial for the couple's preference relation. Hence, with the existence of non-responsive preferences the DA-algorithm cannot be applied. In addition, for some couples markets no stable matching exists (Roth, 1984 and Sotomayor, unpublished note) and the problem of determining if a couples market has a stable matching is NP complete (Ronn, 1990). Unfortunately, this suggests that for non-trivial domains of couples' preferences (*e.g.*, domains containing the subdomain of (weakly) responsive preferences), we might not be able to construct any computationally reasonable algorithm that will either find a stable matching or report an empty output if there is none.

As for the second potential problem, *i.e.*, manipulation by couples acting as single students, we first note that any algorithm that does not accommodate couples is not satisfactory for couples with non-responsive preferences. The new NRMP algorithm does accommodate couples and Roth (2002, p. 1359) explains the reason that single students are processed first as follows: "Because sequencing decisions had no systematic effect on outcomes it was decided to design the algorithm to promote rapid convergence to stability. Based on these computational experiments, the applicant proposing algorithm for the NRMP was designed so that all single applicants are admitted to the algorithm for processing before any couples are admitted." So, it seems that if we want to avoid the problem uncovered in Example 3.2, and we do not mind slower convergence, we may as well switch the priority of the single students over the couples. In that case single students may try to game the mechanism by forming artificial couples with responsive preferences. However, we conjecture that the formation of an artificial couple by two singles is less likely than the artificial split-up of a couple. Alternatively, one may consider to start from scratch and try to construct a completely new algorithm that accommodates couples. Unfortunately, here again Ronn's (1990) NP-completeness result for general couples markets indicates that it might be impossible to find a computationally reasonable algorithm.

⁹In case the preferences are weakly responsive (see Klaus and Klijn, 2005a,b) we can construct another kind of associated individual preferences which enable us again to invoke the DA-algorithm. A couple has weakly responsive preferences if there are individual preferences for the members of the couple that generate, in a similar way as in the case of responsive preferences, the couple's list of *acceptable* combinations of hospitals and unemployment.

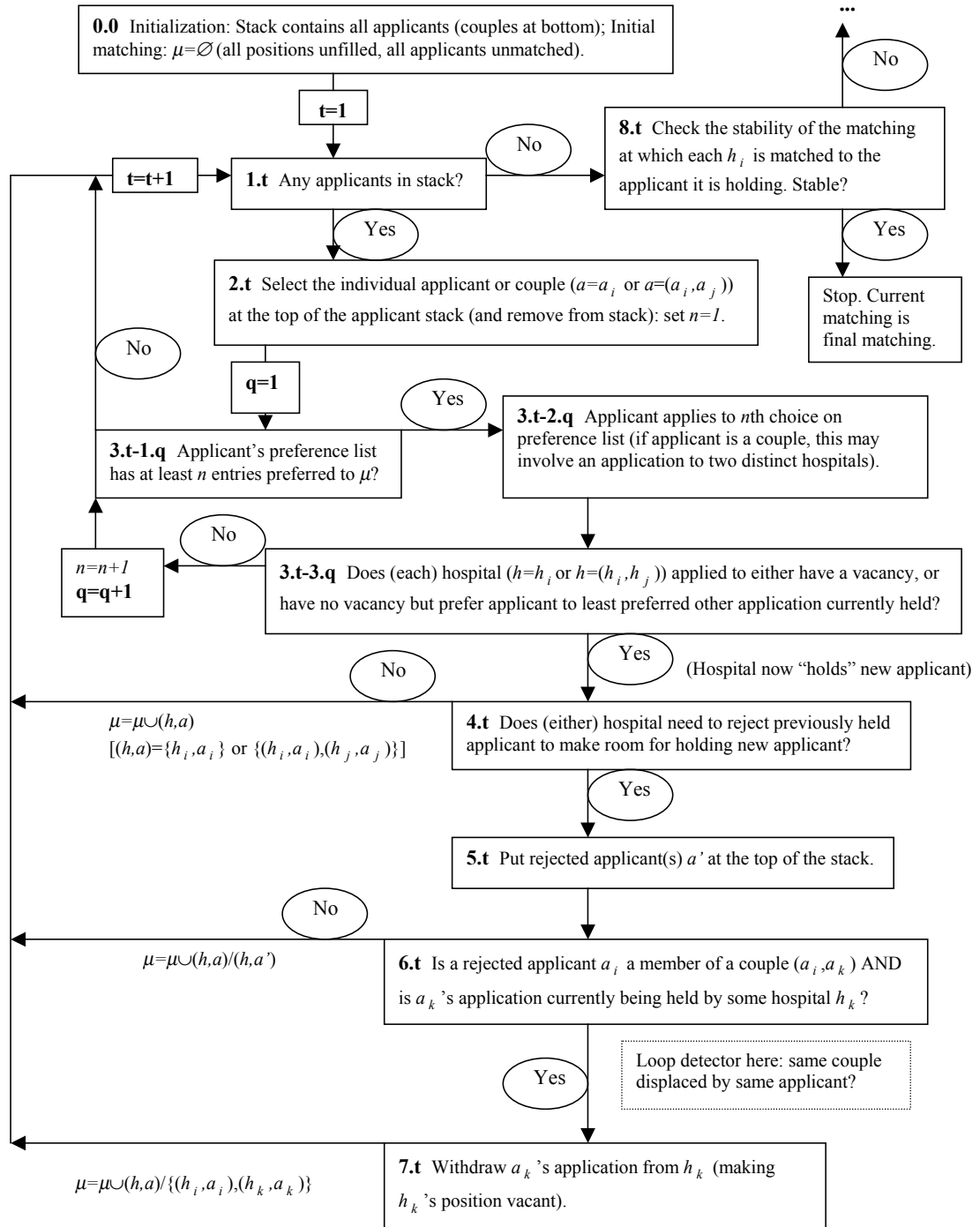
References

- Checker, A. (1973) "The National Intern and Resident Matching Program, 1966-1972," *Journal of Medical Education*, **48**, 107-109.
- Gale, D. and Shapley, L.S. (1962) "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, **69**, 9-15.
- Klaus, B. and Klijn, F. (2005a) "Stable Matchings and Preferences of Couples," *Journal of Economic Theory*, **121**, 75-106.
- Klaus, B. and Klijn, F. (2005b) "Corrigendum: Stable Matchings and Preferences of Couples," UFAE and IAE Working Paper 653-05, Universitat Autònoma de Barcelona.
- Ronn, E. (1990) "NP-Complete Stable Matching Problems," *Journal of Algorithms*, **11**, 285-304.
- Roth, A.E. (1984) "The Evolution of the Labor Market for Medical Interns and Residents: a Case Study in Game Theory," *Journal of Political Economy*, **92**, 991-1016.
- Roth, A.E. (2002) "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics," *Econometrica*, **70**, 1341-1378.
- Roth, A.E. and Peranson, E. (1999) "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, **89**, 748-780.
- Roth, A.E. and Sotomayor, M.A.O. (1990) *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Econometric Society Monograph Series. New York: Cambridge University Press.
- Roth, A.E. and Vande Vate, J.H. (1990) "Random Paths to Stability in Two-Sided Matching," *Econometrica*, **58**, 1475-1480.

5 Appendix

no.	Hospitals				Blocking coalitions?	
	h_1	h_2	h_3	h_4	Students	Hospitals
1	s_1	s_2	s_3	s_4	(s_3, s_4)	(h_1, h_4)
2	s_1	s_2	s_4	s_3	(s_3, s_4)	(h_1, h_3)
3	s_1	s_3	s_2	s_4	(s_1, s_2)	(h_1, h_2)
4	s_1	s_3	s_4	s_2	(s_1, s_2)	(h_1, h_2)
5	s_1	s_4	s_2	s_3	(s_1, s_2)	(h_1, h_2)
6	s_1	s_4	s_3	s_2	(s_1, s_2)	(h_1, h_2)
7	s_2	s_1	s_3	s_4	(s_3, s_4)	(h_2, h_4)
8	s_2	s_1	s_4	s_3	(s_3, s_4)	(h_2, h_3)
9	s_2	s_3	s_1	s_4	—	—
10	s_2	s_3	s_4	s_1	(s_1, s_2)	(h_3, h_1)
11	s_2	s_4	s_1	s_3	(s_1, s_2)	(h_2, h_1)
12	s_2	s_4	s_3	s_1	(s_1, s_2)	(h_2, h_1)
13	s_3	s_1	s_2	s_4	(s_3, s_4)	(h_2, h_4)
14	s_3	s_1	s_4	s_2	(s_1, s_2)	(h_2, h_1)
15	s_3	s_2	s_1	s_4	(s_1, s_2)	(h_3, h_1)
16	s_3	s_2	s_4	s_1	(s_1, s_2)	(h_3, h_1)
17	s_3	s_4	s_1	s_2	(s_1, s_2)	(h_2, h_1)
18	s_3	s_4	s_2	s_1	(s_1, s_2)	(h_2, h_1)
19	s_4	s_1	s_2	s_3	(s_1, s_2)	(h_1, h_2)
20	s_4	s_1	s_3	s_2	(s_1, s_2)	(h_1, h_2)
21	s_4	s_2	s_1	s_3	(s_1, s_2)	(h_1, h_2)
22	s_4	s_2	s_3	s_1	(s_1, s_2)	(h_1, h_2)
23	s_4	s_3	s_1	s_2	(s_1, s_2)	(h_1, h_2)
24	s_4	s_3	s_2	s_1	(s_1, s_2)	(h_1, h_2)

Table 2: Example 3.1, all individually rational matchings (with blocking coalition when possible)



Flowchart 1: The analyzed part of the Applicant Proposing Couples Algorithm (APCA)