# An Integer Programming Model for the Hospitals / Residents Problem with Couples

Iain McBride* and David F. Manlove†

*School of Computing Science, University of Glasgow, UK*
*Email*: i.mcbride.1@research.gla.ac.uk, David.Manlove@glasgow.ac.uk

**Abstract**

The Hospitals / Residents problem with Couples (HRC) is a generalisation of the classical Hospitals / Resident problem (HR) that is important in practical applications because it models the case where couples submit joint preference lists over pairs of (typically geographically close) hospitals. In this paper we give a new NP-completeness result for the problem of deciding whether a stable matching exists, in highly restricted instances of HRC. Further, we present an Integer Programming (IP) model for HRC and extend it the case where preference lists can include ties. Further, we describe an empirical study of an IP model or HRC and its extension to the case where preference lists can include ties. This model was applied to randomly generated instances and also real-world instances arising from previous matching runs of the Scottish Foundation Allocation Scheme, used to allocate junior doctors to hospitals in Scotland.

## 1 Introduction

The National Resident Matching Program (NRMP) matches graduating medical students to hospitals in the US, matching 25,526 students in 2012. Similarly, in Scotland, until recently, medical graduates were matched to Foundation Programme places via the Scottish Foundation Allocation Scheme (SFAS). Centralised matching schemes such as NRMP and SFAS have had to evolve to accommodate linked couples who wish to be allocated to (geographically) compatible hospitals. The requirement to consider the joint preferences of couples has been in place in the NRMP context since 1983 and more recently in the case of SFAS. The underlying allocation problem for NRMP and SFAS can modelled by the so called Hospitals / Residents Problem with Couples (HRC).

An instance of the *Hospitals Residents Problem with Couples* consists of a set of *hospitals* $H$ and a set of *residents* $R$. The residents in $R$ are partitioned into two sets, $S$ and $S'$. The set $S$ consists of *single* residents and the set $S'$ consists of those residents involved in *couples*. There is a set $C = \{(r_i, r_j) : r_i, r_j \in S'\}$ of *couples* such that each resident in $S'$ belongs to exactly one pair in $C$.

Each single resident $r_i \in S$ expresses a linear preference order over some subset of the hospitals in $H$, representing the hospitals that resident $r_i$ finds *acceptable*; any hospital not in this subset is therefore *unacceptable* to $r_i$. Each pair of residents $(r_i, r_j) \in C$ expresses a joint linear preference order over a subset $A$ of $H \times H$ where $(h_p, h_q) \in A$ represents the joint assignment of $r_i$ to $h_p$ and $r_j$ to $h_q$. The hospital pairs in $A$ represent those joint

*Supported by a SICSA Prize PhD Studentship.
†Supported by Engineering and Physical Sciences Research Council grant EP/K010042/1.

assignments that are acceptable to $(r_i, r_j)$, all other joint assignments being unacceptable to $(r_i, r_j)$.

Each hospital $h_j \in H$ expresses a linear preference order over those residents who find $h_j$ acceptable, either as a single resident or as part of a couple. Also, each hospital $h_j \in H$ has a *capacity*, $c_j$, its maximum number of available posts.

The preferences expressed in this fashion are reciprocal: if a resident $r_i$ is acceptable to a hospital $h_j$, either as a single resident or as part of a couple, then $h_j$ is also acceptable to $r_i$, and vice versa. A many-to-one *matching* between residents and hospitals is sought, which is a set of acceptable resident-hospital pairs such that each resident appears in at most one pair and each hospital appears in a number of pairs that does not exceed its capacity. Further, each couple $(r_i, r_j)$ is either jointly unmatched, meaning that both $r_i$ and $r_j$ are unmatched, or jointly matched to some pair $(h_k, h_l)$ that $(r_i, r_j)$ find acceptable.

In an HRC instance we seek a *stable* matching, which guarantees that no resident and hospital, and no couple and pair of hospitals, has an incentive to deviate from their assignments and become matched to each other.

Roth [8] considered stability in the HRC context although did not define the concept explicitly. However, a variety of stability definitions do exist in the HRC context [2, 3, 5]. The definition of stability applied in the work which follows is that given by McDermid and Manlove in [5], shown below in Definition 1, which gives those mutually acceptable pairs, $(r_i, h_k)$ and $((r_i, r_j), (h_k, h_l))$, whose existence would block a matching in HRC.

**Definition 1.** *A matching $M$ is stable if none of the following holds:*

1. *The matching is blocked by a hospital $h_j$ and a single resident $r_i$, as in the classical HR problem.*

2. *The matching is blocked by a couple $(r_i, r_j)$ and a hospital $h_k$ such that* either

   (a) *$(r_i, r_j)$ prefers $(h_k, M(r_j))$ to $(M(r_i), M(r_j))$, and $h_k$ is either undersubscribed in $M$ or prefers $r_i$ to some member of $M(h_k) \backslash \{r_j\}$* or

   (b) *$(r_i, r_j)$ prefers $(M(r_i), h_k)$ to $(M(r_i), M(r_j))$, and $h_k$ is either undersubscribed in $M$ or prefers $r_j$ to some member of $M(h_k) \backslash \{r_i\}$*

3. *The matching is blocked by a couple $(r_i, r_j)$ and (not necessarily distinct) hospitals $h_k \neq M(r_i)$, $h_l \neq M(r_j)$; that is, $(r_i, r_j)$ prefers the joint assignment $(h_k, h_l)$ to $(M(r_i), M(r_j))$, and* either

   (a) *$h_k \neq h_l$, and $h_k$ (respectively $h_l$) is either undersubscribed in $M$ or prefers $r_i$ (respectively $r_j$) to at least one of its assigned residents in $M$;* or

   (b) *$h_k = h_l$, and $h_k$ has at least two free posts in $M$, i.e., $c_k - |M(h_k)| \geq 2$;* or

   (c) *$h_k = h_l$, and $h_k$ has one free post in $M$, i.e., $c_k - |M(h_k)| = 1$, and $h_k$ prefers at least one of $r_i, r_j$ to some member of $M(h_k)$;* or

   (d) *$h_k = h_l$, $h_k$ is full in $M$, $h_k$ prefers $r_i$ to some $r_s \in M(h_k)$, and $h_k$ prefers $r_j$ to some $r_t \in M(h_k) \backslash \{r_s\}$.*

An instance of HRC need not admit a stable matching [9]. Also an instance may admit stable matchings of differing sizes [1]. Further, the problem of deciding whether there exists a stable matching in an instance of HRC is NP-complete, even in the restricted case where there are no single residents and all of the hospitals have only one available post [7, 6].

Let $(\alpha, \beta)$-HRC denote the restriction of HRC in which each single resident's preference list contains at most $\alpha$ hospitals, each couple's preference list contains at most $\alpha$ pairs of hospitals and each hospital's preference list contains at most $\beta$ residents. In many practical applications the residents' preference lists are short. However, the problem remains hard even in this case and Manlove and McDermid [5] showed that $(3, 6)$-HRC is NP-complete.

In Section 2 of this paper we present a new NP-completeness result for the problem of deciding whether there exists a stable matching in an instance of $(2, 3)$-HRC and a summary of an Integer Programming (IP) model for finding a maximum cardinality stable matching in an instance of HRC. Further, in Section 3 we present an empirical study of this model as applied to randomly generated instances and also real-world instances arising from previous matching runs of SFAS. Some conclusions are given in Section 4.

## 2 Complexity of HRC and IP model

In a technical report by the same authors [4] we prove the following new result; for space reasons the details of the proof are omitted.

**Theorem 1.** *Given an instance of* $(2, 3)$-HRC*, the problem of deciding whether the instance supports a stable matching is NP-complete. The result holds even if there are no single residents and each hospital has capacity 1.*

In [4] we give an IP model for finding a maximum cardinality stable matching in HRC. Each model has $O(m)$ binary-valued variables and $O(m + cL^2)$ constraints where $m$ is the total length of the hospitals' preference lists, $c$ is number of couples and $L$ is the maximum length of a couple's preference list. The space complexity of each model is $O(m(m+cL^2))$ and each model can be built in $O(m^4)$ time. For space reasons the details of the models are omitted.

## 3 Empirical Results

We ran experiments on a Java implementation of the IP model as described in [4] applied to both randomly-generated and real data. We present data showing (i) the average time taken to find a maximum cardinality stable matching or report that no stable matching exists, and (ii) the average size of a maximum cardinality stable matching where a stable matching did exist. All experiments were carried out on a desktop PC with an Intel i5-2400 3.1Ghz processor, with 8Gb of memory running Windows 7. The IP solver used in all cases was CPLEX 12.4 and the model was implemented in Java using CPLEX Concert.

To test our implementation for correctness we used a brute force algorithm which recursively generated all possible matchings admitted by an HRC instance and selected a maximum cardinality stable matching from amongst those matchings or reported that none of the generated matchings was stable. Due to the inefficiency of this algorithm it may only be realistically applied to relatively small instances. When solving several thousand HRC instances involving up to 15 residents our implementation agreed with the brute force algorithm when reporting whether the instance admitted a stable solution and further our implementation returned a stable matching of the same size as a maximum cardinality stable matching output by the brute force algorithm.

**Experiments with randomly generated instances.** In our first experiment, we report on data obtained as we increased the number of residents while maintaining a constant ratio of couples, hospitals and posts to residents. For various values of $x$ ($100 \leq x \leq 1000$) in increments of 30, 1,000 randomly generated instances were created containing $x$ residents, $0.1x$ couples and $0.1x$ hospitals with $x$ available posts which were unevenly distributed amongst the hospitals.

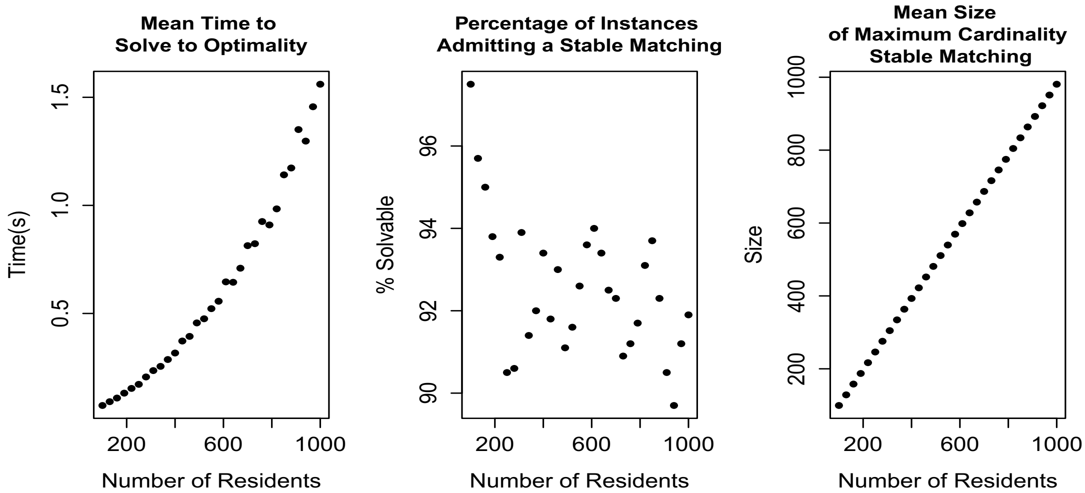


Figure 1: Data obtained when attempting to find a maximum cardinality stable matching in randomly generated instances from Experiment 1.

The data in Figure 1 show that the mean time to find a maximum cardinality stable matching increased as we increased the number of residents in the instance. Figure 1 also shows that the percentage of HRC instances that admit a stable matching does not appear to be correlated with the number of residents involved in the instance and that as the number of residents in the instances increased, the mean size of the maximum cardinality stable matching supported by the instances increased linearly with the number of residents involved the instance.

In the second experiment, we report data as we increased the the percentage of residents involved in couples while maintaining the same total number of residents, hospitals and posts. For various values of $x$ ($0 \leq x \leq 250$) in increments of 25, 1,000 randomly generated instances were created containing 1000 residents, $x$ couples (and hence $1000 - 2x$ single residents) and 100 hospitals with 1000 available posts which were unevenly distributed amongst the hospitals.

The data in Figure 2 show that the mean time to find a maximum cardinality stable matching increased as we increased the number of residents involved in couples. Further, Figure 2 shows that the percentage of HRC instances admitting a stable matching fell as the percentage of residents in the instances involved in couples increased. When 50% of the residents in the instance were involved in a couple, 832 of the 1000 instances admitted a stable matching. Figure 2 also shows that as the percentage of residents in the instances involved in couples increased the mean size of a maximum cardinality stable matching tended to decrease.

**Performance of the model with real world data.** The *Hospitals / Residents Problem with Couples and Ties* (HRCT) is a generalisation of HRC in which hospitals (respectively residents) may find some subsets of their acceptable residents (respectively hospitals) equally preferable. Residents (respectively hospitals) that are found equally preferable by a hospital (respectively resident) are *tied* with each other in the preference list of that hospital (respectively resident). It is straightforward to adapt Definition 1 to the HRCT case.

SFAS assigns junior doctors to two-year training posts in Scotland. In this process the hospitals preferences are derived from the residents' *scores*, where a junior doctor's

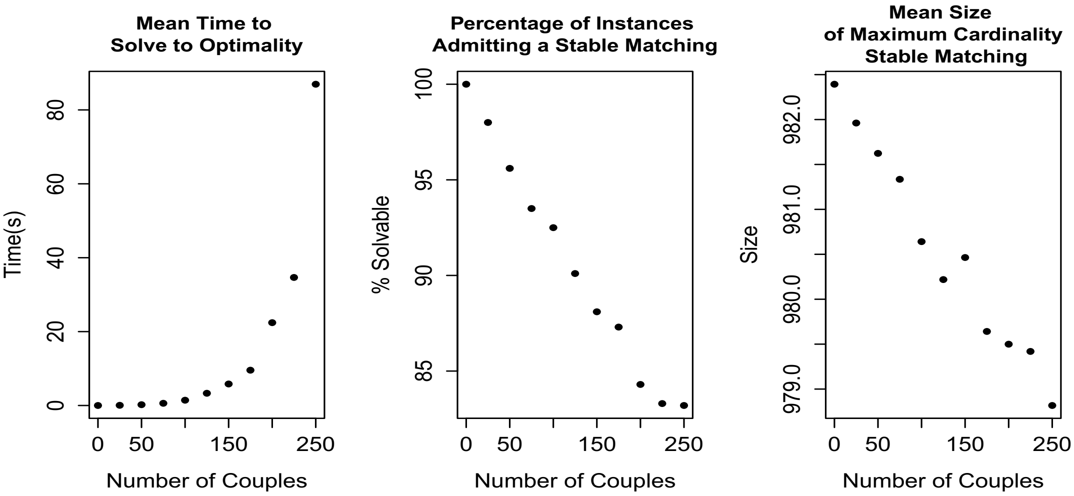


Figure 2: Data obtained when attempting to find a maximum cardinality stable matching in randomly generated instances from Experiment 2.

score is derived from their previous academic performance. If two residents receive the same score, they will be tied in a hospital's preference list. Thus, the underlying SFAS matching problem may be correctly modelled by HRCT.

Hence, we further extended our implementation to solve instances of HRCT as described in [4] and were able to find a maximum cardinality stable matching admitted by the real data obtained from the SFAS context. The sizes of the maximum cardinality stable matchings obtained in the SFAS context for the three years to 2012 are shown in Table 1 alongside the time taken to find these matchings.

## 4 Conclusions

We conclude that the IP model presented in this paper performs well when finding a maximum cardinality stable matching in instances that are similar to those arising in the SFAS application. It remains to investigate the performance of the model as we increase the size of the instance beyond that of the SFAS application.

## References

[1] B. Aldershof and O.M. Carducci. Stable matching with couples. *Discrete Applied Mathematics*, 68:203–207, 1996.

| | Number of Residents | Number of Couples | Number of Hospitals | Number of Posts | Max Cardinality Stable Matching | Time to Solution |
|---|---|---|---|---|---|---|
| 2012 | 710 | 17 | 52 | 720 | 681 | 9.62s |
| 2011 | 736 | 12 | 52 | 736 | 688 | 10.41s |
| 2010 | 734 | 20 | 52 | 735 | 681 | 33.92s |

Table 1: Results obtained from the previous 3 years SFAS data.

[2] P. Biró, R.W. Irving, and I. Schlotter. Stable matching with couples: an empirical study. *ACM Journal of Experimental Algorithmics*, 16, 2011. Section 1, article 2, 27 pages.

[3] D. Gusfield and R.W. Irving. *The Stable Marriage Problem: Structure and Algorithms.* MIT Press, 1989.

[4] I. McBride and D.F Manlove. The Hospitals / Residents Problem with Couples: Complexity and Integer Programming models. Technical report, Computing Research Repository, Cornell University Library, 2013.

[5] E.J. McDermid and D.F. Manlove. Keeping partners together: Algorithmic results for the hospitals / residents problem with couples. *Journal of Combinatorial Optimization*, 19(3):279–303, 2010.

[6] C. Ng and D.S. Hirschberg. Three-dimensional stable matching problems. *SIAM Journal on Discrete Mathematics*, 4:245–252, 1991.

[7] E. Ronn. NP-complete stable matching problems. *Journal of Algorithms*, 11:285–304, 1990.

[8] 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(6):991–1016, 1984.

[9] A.E. Roth. New physicians: A natural experiment in market organization. *Science*, 250:1524–1528, 1990.