The Man-Exchange Stable Marriage Problem

Robert W. Irving

Department of Computing Science, University of Glasgow, Glasgow G12 8QQ, UK.
Email: {rwi}@dcs.gla.ac.uk.

Abstract. We study a variant of the classical stable marriage problem in which there is an additional requirement for a stable matching, namely that there should not be two men each of whom prefers the other's partner in the matching. The problem is motivated by the experience of practical medical matching schemes that use stable matchings, where two students have been known to register a complaint on discovering that each of them would prefer the other's allocation. We show that the problem of deciding whether an instance of the classical stable marriage problem admits a stable matching with this additional man-exchange stability property is NP-complete. This implies a similar result for the hospitals/residents problem, which is a many-to-one generalisation of stable marriage.

Classification Algorithms, computational complexity.

1 Introduction and problem description

The Stable Marriage problem (SM) was introduced in the seminal paper of Gale and Shapley [6]. In its classical form, an instance of SM involves n men and n women, each of whom specifies a *preference list*, which is a total order on the members of the opposite sex. A *matching* M is a set of man-woman pairs such that each person belongs to exactly one pair. If $(m, w) \in M$, we say that w is m's *partner* in M, and vice versa, and we write $M(m) = w$, $M(w) = m$.

We say that a person x *prefers* y to y' if y precedes y' on x's preference list. A matching is *stable* if it admits no *blocking pair*, namely a man m and woman w such that m prefers w to $M(m)$ and w prefers m to $M(w)$. Gale and Shapley proved that every instance of SM admits at least one stable matching; in general, there may be many.

The Hospitals / Residents problem (HR) is a many-to-one generalisation of SM, so called because of its application in centralised matching schemes for the allocation of graduating medical students, or residents, to hospitals [13]. The best known such scheme is the National Resident Matching Program (NRMP) [12] in the US, but similar schemes exist in Canada [5], in Scotland [15, 9], and in a variety of other contexts and countries. In fact, this extension of SM was also discussed by Gale and Shapley under the name of the College Admissions problem. In an instance of HR, each resident has a preference list containing a subset of the hospitals - his/her *acceptable* set, and each hospital ranks the residents for which it is acceptable. In addition, each hospital has a *quota* of available posts, and a matching is a set of resident-hospital pairs so that each resident appears in at most one pair and each hospital in a number of pairs that is bounded by its quota. The definition of stability is easily extended to this more general setting – see, for example, Gusfield and Irving [8] for details. It is again the case that every problem instance admits at least one stable matching. Clearly SM is the special case of HR in which every hospital is acceptable to every resident and each hospital has a quota of 1.

In the stable marriage problem, the stability of a matching M does not preclude the possibility that two men, say m and m', might wish to exchange their partners, since m might prefer $M(m')$ to $M(m)$ and m' might prefer $M(m)$ to $M(m')$. (Of course, the fact that the matching is stable means that, in such a situation, neither of the women involved would prefer to swap.) A similar situation can arise in HR, where two residents might prefer to exchange the hospitals allocated to them in a stable matching, although the hospitals would not sanction such a change. Indeed there is documented evidence [4] of at least one complaint submitted by participants in a medical matching scheme who had discovered just such a circumstance, and who felt that a satisfactory matching should prevent this possibility from arising.

We say that a matching M for an instance of SM is *man-exchange stable* if it is stable in the normal sense, and in addition there do not exist two men m and m' such that m prefers $M(m')$ to $M(m)$ and m' prefers $M(m)$ to $M(m')$. A similar definition of *resident-exchange stability* applies in the HR context. So, in order to pre-empt complaints of the kind mentioned above, a centralised matching scheme might seek to impose a resident-exchange stable matching if one can be found. However, it is easy to see that such a matching need not exist. For example, the stable marriage instance shown in Figure 1 has just two stable matchings, namely $\{(m_1, w_1), (m_2, w_2), (m_3, w_3), (m_4, w_4)\}$ and $\{(m_1, w_2), (m_2, w_4), (m_3, w_3), (m_4, w_1)\}$. The first of these has the *man-exchange blocking pair* $\{m_2, m_3\}$ while the second is blocked by the pairs $\{m_1, m_2\}$ and $\{m_1, m_4\}$.

Men's preferences					Women's preferences				
m_1 :	w_1	w_4	w_2	w_3	w_1 :	m_4	m_1	m_2	m_3
m_2 :	w_3	w_2	w_4	w_1	w_2 :	m_1	m_2	m_4	m_3
m_3 :	w_2	w_3	w_4	w_1	w_3 :	m_3	m_2	m_1	m_4
m_4 :	w_3	w_2	w_4	w_1	w_4 :	m_2	m_4	m_3	m_1

Figure 1: A stable marriage instance with no man-exchange stable matching.

On the other hand, the instance shown in Figure 2 has just one stable matching, namely $\{(m_1, w_4), (m_2, w_2), (m_3, w_1), (m_4, w_3)\}$, and it is straightforward to check that this is also man-exchange stable.

Men's preferences					Women's preferences				
m_1 :	w_1	w_4	w_3	w_2	w_1 :	m_2	m_3	m_1	m_4
m_2 :	w_4	w_2	w_1	w_3	w_2 :	m_2	m_4	m_1	m_3
m_3 :	w_2	w_1	w_4	w_3	w_3 :	m_1	m_4	m_3	m_2
m_4 :	w_2	w_3	w_4	w_1	w_4 :	m_1	m_4	m_2	m_3

Figure 2: A stable marriage instance with a man-exchange stable matching.

Unfortunately, there is little prospect of an efficient algorithm to determine, for a given instance of SM whether there exists a man-exchange stable matching (or one that is resident-exchange stable for an instance of HR), and if so to find one – it is the main purpose of this paper to show that the decision problem is NP-complete.

The formal description of the Man-Exchange Stable Marriage problem is as follows.

Man-Exchange Stable Marriage (MESM)
INSTANCE: A set of n men and a set of n women, and for each person a total order on

the persons of the opposite sex.
QUESTION: Does there exist a one-to-one matching of the men to the women that (a) is stable in the classical sense, and (b) admits no man-exchange blocking pair.

A similar concept of exchange-stability was introduced by Alcalde [1] in the context of the non-bipartite *stable roommates* problem, but without the pre-requisite that the matching be stable in the classical sense. A version of this problem was subsequently shown to be NP-complete by Cechlárová [2]. Cechlárová and Manlove [3] studied a number of variants of exchange-stable matching. They showed that, if the requirement for stability in the classical sense is dropped, then a matching that is man-exchange stable in this weaker sense is guaranteed to exist, and can be found in polynomial time. On the other hand, if, in addition, we allow a pair of women, as well as a pair of men, to form a blocking pair, (but still do not require classical stability) then the resulting Exchange-stable Marriage problem is NP-complete.

However, none of these results has implications for our notion of Man-Exchange Stability, where the pre-requisite of stability in the classical sense completely changes the nature of the problem. Indeed, the existence of a polynomial-time algorithm for MESM was posed as an open problem in [3]. (Note that, in [3], the term 'Man-exchange stable marriage' is used to refer to the version of the problem where classical stability is not required; our use of terminology differs from this.)

The remainder of this paper is structured as follows. Section 2 summarises the Gale-Shapley algorithm for the stable marriage problem, and the concepts of man-optimal and woman-optimal stable matchings. In Section 3, we review the key results that characterise the lattice structure of the set of stable matchings. These are needed to understand the ideas underlying the NP-completeness reduction. This reduction, using as starting point a variant of Satisfiability, is presented in Section 4, together with the necessary correctness proofs. We conclude with a summary and some additional observations in Section 5.

2 The Gale-Shapley algorithm

The (man-oriented version of the) Gale-Shapley algorithm involves a sequence of 'proposals' from men to women. Each proposal may be accepted – if the woman is free, or if she prefers this man to her current partner – and otherwise is rejected; if it is accepted, the man and woman become partners, and she rejects her current partner, if any, at that point, setting him free. This process continues until there is no free man, at which point the set of partners forms a stable matching. A version of the Gale-Shapley algorithm is presented in Figure 3.

```
assign each person to be free;
while (some man m is free) {
    w = first woman on m's list to whom he has not proposed;
    if (w is free)
        w accepts m;
    elsif (w prefers m to her partner m')
        w rejects m' and accepts m;
    else
        w rejects m;
}
```

Figure 3: The Gale-Shapley Algorithm

It is well known that the apparent non-determinism inherent in the algorithm is immaterial, and that all possible executions of the Gale-Shapley algorithm find the *man-optimal* stable matching, i.e., the stable matching in which every man has the best partner that he can have in any stable matching (and, as it turns out, every woman the worst). Of course, interchanging the roles of men and women in the algorithm leads to the *woman-optimal* stable matching, which is simultaneously optimal for all of the women, and 'pessimal' for all of the men. For further details, see the comprehensive treatment of Gusfield and Irving [8].

In some sense the man-optimal stable matching has the best chance of being man-exchange stable, since the further down a man's preference list that his partner is, the more possibilities that man has to be a member of a man-exchange blocking pair. However, it is easy to see that there are cases where the man-optimal is not exchange-stable but some other stable matching is – it may even be the case that there are many stable matchings, but only the woman-optimal is man-exchange stable.

For example, the instance shown in Figure 4 has 3 stable matchings. The man-optimal $\{(m_1, w_1), (m_2, w_2), (m_3, w_3), (m_4, w_4), (m_5, w_5), (m_6, w_6)\}$ is blocked by the men $\{m_1, m_2\}$, the stable matching $\{(m_1, w_1), (m_2, w_2), (m_3, w_4), (m_4, w_3), (m_5, w_5), (m_6, w_6)\}$ is blocked by the men $\{m_3, m_4\}$, while $\{(m_1, w_3), (m_2, w_4), (m_3, w_5), (m_4, w_6), (m_5, w_1), (m_6, w_2)\}$, which is the woman-optimal, is man-exchange stable.

Men's preferences							Women's preferences						
m_1 :	w_2	w_1	w_3	w_4	w_6	w_5	w_1 :	m_5	m_1	m_6	m_2	m_3	m_4
m_2 :	w_1	w_2	w_3	w_4	w_5	w_6	w_2 :	m_6	m_2	m_1	m_4	m_3	m_5
m_3 :	w_3	w_4	w_2	w_5	w_1	w_6	w_3 :	m_1	m_4	m_3	m_6	m_2	m_5
m_4 :	w_4	w_1	w_3	w_5	w_6	w_2	w_4 :	m_2	m_3	m_4	m_5	m_1	m_6
m_5 :	w_5	w_1	w_4	w_6	w_3	w_2	w_5 :	m_3	m_5	m_6	m_4	m_2	m_1
m_6 :	w_1	w_6	w_2	w_3	w_5	w_4	w_6 :	m_4	m_6	m_3	m_1	m_5	m_2

Figure 4: The woman-optimal can be the only man-exchange stable matching.

3 Rotations and the rotation poset

For a given stable marriage instance, we define a partial order relation on the set of stable matchings as follows: matching M *dominates* matching M' if every man has at least as good a partner in M as he has in M'. It turns out that the set of stable matchings forms a distributive lattice under this partial order [11]; in the *meet* $M \wedge M'$, each man has the better of his partners in M and M', and in the *join* $M \vee M'$, each man has the poorer of these two partners (or, in each case, the same partner as in M and M' if these coincide). The man-optimal stable matching M_m and the *woman-optimal* M_w are the top and bottom of this lattice, respectively.

An instance of SM can have exponentially many stable matchings [11]. However, the so-called *rotation poset*, first described by Irving and Leather [10], gives a compact representation of the lattice. We will exploit this structural characterisation in our NP-completeness proof in Section 4.

Let M be a stable matching for an instance I of SM. For each man m such that $M(m) \neq M_w(m)$, we define the *successor* woman for m relative to M, denoted $s_M(m)$ to be the first woman w on m's list, following $M(m)$, such that w prefers m to $M(w)$. There must be such a woman if $M(m) \neq M_w(m)$, since $M_w(m)$ qualifies, but $s_M(m)$ is undefined if $M(m) = M_w(m)$.)

A *rotation exposed* in M is a cyclic sequence of pairs $\rho = (m_0, w_0), \ldots, (m_{r-1}, w_{r-1})$ such that $M(m_i) = w_i$ and $s_M(m_i) = w_{i+1}$ for all i, $i+1$ taken modulo r. To *eliminate* the exposed rotation ρ from the stable matching M means to pair m_i with w_{i+1} for all m_i in the rotation ($i+1$ again taken modulo r) and otherwise leave the pairs unchanged. We denote the resulting matching, which also turns out to be stable, by M/ρ. Every stable matching $M \neq M_w$ has at least one exposed rotation; to find one, choose a man x_0 for whom $M(x_0) \neq M_w(x_0)$ and generate the sequence $x_{i+1} = M(s_M(x_i))$ until it cycles, setting $m_i = x_{t+i}$, where x_t is the first repeated man in the sequence. It turns out that every stable matching can be obtained by starting with the man-optimal and successively eliminating a sequence of exposed rotations. Each stable matching is characterised by the set of rotations that must be eliminated to produce it.

A partial order relation can be defined on the rotations for an SM instance as follows: rotation ρ precedes rotation ρ', written $\rho \prec \rho'$, if, in order to obtain a stable matching in which ρ' is exposed, rotation ρ must be eliminated. The set of rotations under this partial order is known as the *rotation poset*. A subset S of the rotation poset with the property that $\rho \in S, \sigma \prec \rho \Rightarrow \sigma \in S$ is called *closed*. The following theorem gives what, for our purposes, is the key characterisation of the set of stable matchings in terms of the rotation poset.

Theorem 1 *For an instance of Stable Marriage, there is a one-to-one correspondence between the stable matchings and the closed subsets of the rotation poset. A stable matching can be obtained from the man-optimal by eliminating, in any valid sequential order, the rotations in the corresponding closed subset.*

We refer to the unique closed subset of the rotation poset corresponding to a given stable matching M simply as the *rotation set* of M. Full coverage of the structural aspects of stable marriage, and how these can be expoited to give efficient algorithms for a range of problems related to the set of stable matchings, is given by Gusfield and Irving [8].

4 NP-completeness of man-exchange stable marriage

We prove that man-exchange stable marriage is NP-complete by describing a polynomial-time reduction from a variant of Satisfiability. This variant is defined as follows:

Unnegated One-in-Three 3SAT (U-1in3-3SAT)
INSTANCE: A set U of variables, a collection $\mathcal{C}$ of clauses over U such that each clause $C \in \mathcal{C}$ has $|C| = 3$ and no clause contains a negated variable.
QUESTION: Is there a truth assignment for U such that each clause in $\mathcal{C}$ has exactly one true literal.

This problem was shown to be NP-complete by Schaefer [14] – see also [7, Problem LO4, Page 259].

First of all we describe how, given an instance of U-1in3-3SAT, we construct an instance of MESM, and then, via a sequence of lemmas we show that this construction has all of the properties required of a polynomial-time reduction.

Let I be an instance of U-1in3-3SAT involving n variables $a_1, \ldots, a_n$ and t clauses $C_1, \ldots, C_t$. We introduce $5tn$ men m_{ijk} and $5tn$ women w_{ijk} ($1 \leq i \leq n, 1 \leq j \leq t, 1 \leq k \leq 5$), and an additional n men x_i and n women y_i ($1 \leq i \leq n$).

The preference lists for the men are defined as follows:

If $a_i \in C_j = \{a_i, a_p, a_q\}$

$$\begin{array}{lllllllll}
m_{i11}: & A & w_{i11} & B & w_{i12} & C & w_{i13} & \ldots & \\
m_{ij1}: & & w_{ij1} & B & w_{ij2} & C & w_{ij3} & \ldots & (j>1) \\
m_{ij2}: & & w_{ij2} & & w_{ij3} & & w_{ij4} & \ldots & \\
m_{ij3}: & & w_{ij3} & & w_{ij4} & & w_{ij5} & \ldots & \\
m_{ij4}: & & w_{ij4} & & w_{ij5} & & w_{i(j+1)1} & \ldots & \\
m_{ij5}: & & w_{ij5} & & w_{i(j+1)1} & & w_{ij2} & \ldots &
\end{array}$$

If $a_i \notin C_j$

$$\begin{array}{llllll}
m_{i11}: & A & w_{i11} & w_{i12} & \ldots & \\
m_{ij1}: & & w_{ij1} & w_{ij2} & \ldots & (j>1) \\
m_{ij2}: & & w_{ij2} & w_{ij3} & \ldots & \\
m_{ij3}: & & w_{ij3} & w_{ij4} & \ldots & \\
m_{ij4}: & & w_{ij4} & w_{ij5} & \ldots & \\
m_{ij5}: & & w_{ij5} & w_{i(j+1)1} & \ldots &
\end{array}$$

where

$$\begin{array}{lll}
A & = & \{w_{s11} : a_i \text{ and } a_s \text{ appear together in some clause}\} \\
B & = & \{w_{pj2}, w_{qj2}\} \\
C & = & \{w_{pj3}, w_{qj3}\}.
\end{array}$$

Here, a set of women included at a certain position in a preference list may appear in any order, the symbol $\ldots$ indicates all remaining women not explicitly listed earlier, and $(j+1)$ becomes 1 in the case where $j = t$.

Finally,

$$x_i : \quad w_{i11} \quad y_i \quad \ldots$$

The preference lists for the women are defined as follows:

$$\begin{array}{llllll}
w_{i11}: & m_{it4} & m_{it5} & m_{i11} & x_i & \ldots \\
w_{ij1}: & m_{i(j-1)4} & m_{i(j-1)5} & m_{ij1} & \ldots & (j>1) \\
w_{ij2}: & m_{ij5} & m_{ij1} & m_{ij2} & \ldots & \\
w_{ij3}: & m_{ij1} & m_{ij2} & m_{ij3} & \ldots & \\
w_{ij4}: & m_{ij2} & m_{ij3} & m_{ij4} & \ldots & \\
w_{ij5}: & m_{ij3} & m_{ij4} & m_{ij5} & \ldots & \\
y_i: & x_i & \ldots & & &
\end{array}$$

Again, the symbol $\ldots$ indicates all remaining men not explicitly listed earlier.

Before presenting the formal proofs, we give an overview of the rationale for this construction.

The rotation structure for the derived instance of MESM is as follows:

- for each i $(1 \leq i \leq n)$

  $$\rho_i = ((m_{i11}, w_{i11}), (m_{i12}, w_{i12}), \ldots, (m_{i15}, w_{i15}), (m_{i21}, w_{i21}), \ldots, (m_{it5}, w_{it5}))$$

  is a rotation of length $5t$ that has no predecessors in the rotation poset;

- for each i, j $(1 \leq i \leq n, 1 \leq j \leq t)$ such that $a_i \in C_j$,

  $$\sigma_{ij} = ((m_{ij1}, w_{ij2}), (m_{ij2}, w_{ij3}), (m_{ij3}, w_{ij4}), (m_{ij4}, w_{ij5}), (m_{ij5}, w_{i(j+1)1}))$$

  is a rotation of length 5 that is a successor of ρ_i, and only ρ_i, in the rotation poset.

The entries A in the men's preference lists ensure that, for each clause $C_j = \{a_p, a_q, a_r\}$, at least two of the rotations ρ_p, ρ_q and ρ_r must be in a rotation set that corresponds to a stable matching that avoids a man-exchange blocking pair. Furthermore, the entries B and C ensure that, if all three of these rotations are in a rotation set then a man-exchange blocking pair for the corresponding stable matching cannot be avoided. Hence we can have a man-exchange stable matching only if we can choose a subset of the original variables that has an intersection of size 2 with each clause - and this subset represents the false variables in a satisfying assignment. The argument works in the other direction too.

We now formalise this argument, and thereby establish the validity of our reduction.

Lemma 2 *For the derived instance of MESM, the man-optimal stable matching pairs m_{ijk} with w_{ijk} for all i, j, k, and x_i with y_i for all i.*

Proof: As described in Section 2, the man-optimal stable matching is produced by applying the Gale-Shapley algorithm from the men's side, and in so doing the order in which men propose is immaterial. A possible execution of the algorithm is as follows:

1. m_{ij2} proposes to w_{ij2} for all i, j; all proposals accepted;
2. m_{ij3} proposes to w_{ij3} for all i, j; all proposals accepted;
3. m_{ij4} proposes to w_{ij4} for all i, j; all proposals accepted;
4. m_{ij5} proposes to w_{ij5} for all i, j; all proposals accepted;
5. x_i proposes to w_{i11} for all i; all proposals accepted;
6. m_{i11} proposes to all the women in set A for all i; all proposals rejected because each woman in A prefers the man x who proposed to her previously;
7. m_{ij1} proposes to w_{ij1} for all i and j; all proposals accepted, and all x_i rejected;
8. x_i proposes to y_i for all i; all proposals accepted.

The algorithm terminates at this point, and it is clear that the matching produced is the one listed in the statement of the lemma, which is therefore the man-optimal stable matching. ■

Lemma 3 *For the derived instance of MESM, each of the cyclic sequences ρ_i defined above is a rotation that is exposed in the man-optimal M_m, and which therefore has no predecessors in the rotation poset.*

Proof: It is immediate from the previous lemma that $M_m(m_{ijk}) = w_{ijk}$ for all j and k. By inspection of the preference lists, it is also immediate that $s_{M_m}(m_{ij2}) = w_{ij3}$, $s_{M_m}(m_{ij3}) = w_{ij4}$, $s_{M_m}(m_{ij4}) = w_{ij5}$, and $s_{M_m}(m_{ij5}) = w_{i(j+1)1}$ for all j, where $j + 1$ becomes 1 in the case $j = t$. Furthermore, in the case of m_{ij1}, all of the women in the set B prefer their partner in M_m to m_{ij1}, from which it follows, again by inspection of the preference lists, that $s_{M_m}(m_{ij1}) = w_{ij2}$ for all j. ■

Lemma 4 *For the derived instance of MESM, each of the cyclic sequences σ_{ij} defined above is a rotation that has only ρ_i as a predecessor in the rotation poset.*

Proof: Consider the stable matching M_i obtained from M_m by the elimination of rotation ρ_i. It is immediate that $M_i(m_{ij1}) = w_{ij2}$, $M_i(m_{ij2}) = w_{ij3}$, $M_i(m_{ij3}) = w_{ij4}$, $M_i(m_{ij4}) = w_{ij5}$, and $M_i(m_{ij5}) = w_{i(j+1)1}$ for all relevant j. By inspection of the preference lists, it is also immediate that $s_{M_i}(m_{ij2}) = w_{ij4}$, $s_{M_i}(m_{ij3}) = w_{ij5}$, $s_{M_i}(m_{ij4}) = w_{i(j+1)1}$, and $s_{M_i}(m_{ij5}) = w_{ij2}$ for all relevant j. In the case of m_{ij1}, all of the women in the set C prefer their partner in M_i to m_{ij1}, from which it follows, again by inspection of the preference lists, that $s_{M_i}(m_{ij1}) = w_{ij3}$ for all relevant j.

Hence σ_{ij} is a rotation exposed in M_i. But this rotation clearly cannot be exposed unless ρ_i has been eliminated, so it has ρ_i, and only ρ_i as a predecessor. ■

Lemma 5 *Suppose that the derived instance of MESM admits a man-exchange stable matching M, and let S be the rotation set corresponding to M. Then for each clause $C_j = \{a_p, a_q, a_r\}$, exactly two of ρ_p, ρ_q, ρ_r are in S.*

Proof: Consider the man-optimal stable matching M_m. From the preference lists we can verify that $\{m_{p11}, m_{q11}\}$ forms a man-exchange blocking pair, since m_{p11} prefers w_{q11} (who is in the relevant set A) to w_{p11}, and likewise m_{q11} prefers w_{p11} to w_{q11}. Similarly $\{m_{p11}, m_{r11}\}$ and $\{m_{q11}, m_{r11}\}$ also form man-exchange blocking pairs. Since the pairs (m_{p11}, w_{p11}), (m_{q11}, w_{q11}) and (m_{r11}, w_{r11}) are in the rotations ρ_p, ρ_q and ρ_r respectively, it follows that at least two of these rotations must be eliminated to avoid these three man-exchange blocking pairs.

Suppose that ρ_p and ρ_q are in S. If neither σ_{pj} nor σ_{qj} is in S then $M(m_{pj1}) = w_{pj2}$ and $M(m_{qj1}) = w_{qj2}$, so that $\{m_{pj1}, m_{qj1}\}$ forms a man-exchange blocking pair for M because of the women in B. On the other hand, if both σ_{pj} and σ_{qj} are in S then $M(m_{pj1}) = w_{pj3}$ and $M(m_{qj1}) = w_{qj3}$, so that this same pair again forms a man-exchange blocking pair for M, this time because of the women in C. If all three of ρ_p, ρ_q, ρ_r are in S, then we would require exactly one of σ_{pj}, σ_{qj}, exactly one of σ_{pj}, σ_{rj}, and exactly one of σ_{qj}, σ_{rj} to be in S, which is clearly impossible. It follows that exactly two of ρ_p, ρ_q, ρ_r are in S. ■

Lemma 6 *Suppose that the derived instance of MESM admits a man-exchange stable matching M. Then the given instance of U-1in3-3SAT admits a satisfying assignment.*

Proof: From Lemma 5(i), it follows that the closed subset S of the rotation poset corresponding to a stable matching that is also man-exchange stable must contain exactly two of the rotations ρ_i corresponding to each clause in the given instance. Consequently, if we assign variable a_i to be false if and only if $\rho_i \in S$, the resulting assignment will satisfy the given boolean expression with exactly one true literal in each clause, as required. ■

Lemma 7 *Suppose that the given instance of U-1in3-3SAT admits a satisfying assignment. Then the derived instance of MESM admits a man-exchange stable matching.*

Proof: Given a satisfying assignment for the instance of U-1in3-3SAT, define the subset S of the rotation poset to consist of all of the rotations ρ_i corresponding to variables a_i that are assigned false, and for each clause $C_j = \{a_p, a_q, a_r\}$ in which, say a_p and a_q are assigned false, exactly one of the rotations σ_{pj}, σ_{qj}. It follows from Lemmas 3 and 4 that S is a closed subset of the rotation poset. We claim that the stable matching M corresponding to set S is man-exchange stable.

Firstly, suppose that a_i is a true variable. Man m_{i11} is paired with w_{i11} in M, so prefers only women w_{s11} such that variable a_s is in some clause with a_i. So a_s must be a false variable, and w_{s11} is therefore paired with either m_{st5} or m_{st4} in M. But neither of these men prefers w_{i11} to w_{s11}, so m_{i11} cannot be a member of a man-exchange blocking pair. Any man m_{ijk} ($j \neq 1$ or $k \neq 1$) has his first choice woman as partner in M, so also cannot be a member of a man-exchange blocking pair.

Secondly, if a_i is a false variable, consider whether any man m_{ij2} could be in a man-exchange blocking pair. Such a man prefers only w_{ij2}, and possibly w_{ij3} to his partner in M. But the partner of w_{ij2} is either m_{ij1} or m_{ij5}, and neither of these men prefers m_{ij2}'s partner to his own. If m_{ij2} prefers w_{ij3} to his partner w_{ij4}, then w_{ij3} must have m_{ij1} as partner, and he does not prefer w_{ij4}. The same argument holds for m_{ij3} and m_{ij4}. In the case of m_{ij5}, he prefers only w_{ij5}, and possibly $w_{i(j+1)1}$ to his partner in M, and once again it is easy to check that neither of these women's partners prefers m_{ij5}'s partner to his own.

Next, for a false variable a_i, consider whether a man m_{ij1} could be in a man-exchange blocking pair. Suppose first that $\sigma_{ij} \notin S$, so that $M(m_{ij1}) = w_{ij2}$. Then m_{ij1} prefers the following women to his partner in M:

1. w_{ij1};
2. w_{pj2} and w_{qj2}, where $C_j = \{a_i, a_p, a_q\}$;
3. when $j = 1$, women w_{s11} for all values of s such that a_i and a_s occur together in some clause.

In case (1), $M(w_{ij1}) = m_{i(j-1)r}$ for $r = 4$ or $r = 5$, and these men prefer their own partners to w_{ij2}. For case (2), we know that one of a_p, a_q is true and the other false; suppose that a_p is true. Then $\rho_p \notin S$, so that m_{pj2} has his first choice woman w_{pj2} as his partner in M. It follows that a_q is false, so $\rho_q \in S$ and further we must have $\sigma_{qj} \in S$, so $M(w_{qj2}) = m_{qj5}$. But inspection of m_{qj5}'s preference list reveals that he does not prefer w_{ij2} to his own partner. We subdivide case (3) into two subcases. If $\rho_s \notin S$ then $M(w_{s11}) = m_{s11}$, and m_{s11} does not prefer w_{i12}. If $\rho_s \in S$ then $M(w_{s111}) = m_{st4}$ or m_{st5}, and again neither of these men prefers w_{i12}.

Finally, suppose that $\sigma_{ij} \in S$, so that $M(m_{ij1}) = w_{ij3}$. Then m_{ij1} prefers the following women to his partner in M:

1. w_{ij1};
2. w_{ij2};
3. w_{pj2} and w_{qj2}, where $C_j = \{a_i, a_p, a_q\}$;
4. w_{pj3} and w_{qj3}, where $C_j = \{a_i, a_p, a_q\}$;
5. when $j = 1$, women w_{s11} for all values of s such that a_i and a_s occur together in some clause.

In case (1), $M(w_{ij1}) = m_{i(j-1)r}$ for $r = 4$ or $r = 5$, and these men prefer their own partners to w_{ij3}. In case (2), $M(w_{ij2}) = m_{ij5}$, and this man prefers his own partner w_{ij2} to w_{ij3}.

For cases (3) and (4),again assume that a_p is true and a_q false, so that $\rho_p \notin S$ and $\rho_q \in S$. In case (3), $M(w_{pj2}) = m_{pj2}$ and m_{pj2} prefers w_{pj2} to w_{ij3}, while $M(w_{qj2}) = m_{qj1}$, since $\sigma_{qj} \notin S$, and m_{qj1} prefers w_{qj2} to w_{ij3}. In case (4), $M(w_{pj3}) = m_{pj3}$ and m_{pj3} prefers w_{pj3} to w_{ij3}, while $M(w_{qj3}) = m_{qj2}$, since $\sigma_{qj} \notin S$, and m_{qj2} prefers w_{qj3} to w_{ij3}.

Lastly, we subdivide case (5) into two subcases. If $\rho_s \notin S$ then $M(w_{s11}) = m_{s11}$, and m_{s11} does not prefer w_{i13}. If $\rho_s \in S$ then $M(w_{s111}) = m_{st4}$ or m_{st5}, and again neither of these men prefers w_{i13}. ■

Theorem 8 *The man-exchange stable marriage problem is NP-complete.*

Proof: Membership in NP is immediate. NP-completeness follows from Lemmas 6 and 7, together with the known NP-completeness of U-1in3-3SAT. ∎

A more constrained version of the MESM problem arises if we also insist that there should be no two women each of whom prefers the other's partner to her own; we refer to the problem of determining whether there is a stable matching with this even stronger exchange-stability property as General Exchange Stable Marriage (GESM). Perhaps it is not surprising that this variant also turns out to be NP-complete.

Theorem 9 *The general exchange stable marriage problem is NP-complete.*

Proof: We use the same transformation as for MESM. Examination of the women's preference lists, and the set of stable matchings, for the derived instance, reveals that none of the stable matchings has an exchange blocking pair of women, and the result follows at once from this observation. ∎

5 Summary and conclusion

In this paper we have introduced the Man-Exchange Stable Marriage problem, motivated by a significant practical application, and have shown that this problem, and therefore its generalisation in the Hospitals / Residents context, is NP-complete.

It is natural to ask how the situation is affected if we allow a stable matching to be blocked by any coalition of $t \geq 2$ men, all of whom could improve if they permuted their partners in a suitable way. In the context of practical matching schemes, this *man-coalition stable marriage problem* could still be significant although, in reality, the chances that a set of more than two residents would collectively realise that they could form such a coalition would be much diminished.

In fact, for this extended problem, the situation is rather different, as the following theorem, first established by Cechlarova and Manlove [3] indicates.

Theorem 10 *For any stable marriage instance, the man-optimal stable matching is the only stable matching that can be man-coalition stable.*

Proof: Let M be any stable matching other than the man-optimal, and let S be the corresponding (non-empty) rotation set. Let $\rho = (m_0, w_0), \ldots, (m_{r-1}, w_{r-1})$ be a member of S with no successors in S. Then, for each i ($0 \leq i \leq r-1$), m_i prefers w_i to w_{i+1}, his partner in M ($i+1$ taken modulo r). Hence the set $\{m_0, \ldots m_{r-1}\}$ is a blocking coalition for M. ∎

It is a simple matter to determine, in $O(n^2)$ time whether a stable marriage instance involving n men admits a man-coalition stable matching. Construct a directed graph G with a vertex v_i for each man m_i and a directed edge from v_i to v_j if and only if m_i prefers $M(m_j)$ to $M(m_i)$, where M is the man-optimal stable matching. Then it is easy to show that M is man-coalition stable if and only if G is acyclic. The man-optimal stable matching can be found in $O(n^2)$ time by the Gale-Shapley algorithm, and the existence of a cycle in G can be checked also in $O(n^2)$ time by standard means such as depth-first search.

References

[1] J. Alcalde. Exchange-proofness or divorce-proofness? Stability in one-sided matching markets. *Economic Design*, 1:275–287, 1995.

[2] K. Cechlárová. On the complexity of exchange stable roommates. *Discrete Applied Mathematics*, 116:279-287, 2002.

[3] K. Cechlárová and D.F. Manlove. The exchange stable marriage problem. Technical Report TR2003-142, Computing Science Department, University of Glasgow, September 2003.

[4] J. Allan, NHS Scotland. Personal communication, 2002.

[5] Carms: the Canadian Resident Matching Scheme. Web document available at `http://www.carms.ca/main.htm`.

[6] D. Gale and L.S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69:9–15, 1962.

[7] M.J Garey and D.S. Johnson. *Computers and Intractability: a Guide to the Theory of NP-completeness.* Freeman, 1979.

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

[9] R. W. Irving. Matching medical students to pairs of hospitals: a new variation on an old theme. *Proceedings of ESA'98, the Sixth Annual European Symposium on Algorithms*, Venice Italy, 1998, Lecture Notes in Computer Science, vol. 1461:381-392, Springer 1998.

[10] R.W. Irving and P. Leather. The complexity of counting stable marriages. *SIAM Journal on Computing*, 15:655–667, 1986.

[11] D.E. Knuth. *Mariages Stables.* Les Presses du l'Universitè de Montrèal, 1976.

[12] National Resident Matching Program. About the NRMP. Web document available at `http://www.nrmp.org/about_nrmp/how.html`.

[13] 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.

[14] T.J. Schaefer. The complexity of satisfiability problems. *Proc. 10th. Ann. ACM Symp. on Theory of Computing*, ACM New York, 216–226, 1978.

[15] SPA: Scottish Pre-registration House Officer Allocations. Web document available at `http://www.nes.scot.nhs.uk/spa/front.htm`.