Title
Lower bounds for the stable marriage problem and its variants

Authors
Ng, Cheng
Hirschberg, Daniel S.

Publication Date
1988

Peer reviewed

Lower Bounds for the Stable Marriage Problem and its Variants

Cheng Ng and Daniel S. Hirschberg

Technical Report No. 88-18

Abstract

In an instance of the stable marriage problem of size n, n men and n women each ranks members of the opposite sex in order of preference. A *stable marriage* is a complete matching $M = \{(m_1, w_{i_1}), (m_2, w_{i_2}), \ldots, (m_n, w_{i_n})\}$ such that no unmatched man and woman prefer each other to their partners in M.

A pair (m_i, w_j) is *stable* if it is contained in some stable marriage. In this paper, we prove that determining if an arbitrary pair is stable requires $\Omega(n^2)$ time in the worst case. We show, by an adversary argument, that there exists instances of the stable marriage problem such that it is possible to find at least one pair that exhibits the $\Omega(n^2)$ lower bound.

As corollaries of our results, the lower bound of $\Omega(n^2)$ is established for several stable marriage related problems. Knuth, in his treatise on stable marriage, asks if there is an algorithm that finds a stable marriage in less than $\Theta(n^2)$ time. Our results show that such an algorithm does not exist.

Lower Bounds for the Stable Marriage Problem and its Variants

Introduction

An instance of the stable marriage problem involves two disjoint sets of equal cardinality n, the men denoted by m_i's and women denoted by w_i's. Each individual ranks all members of the opposite sex in order of decreasing preference. It is useful to represent these preferences in two $n \times n$ integer matrices MP and WP such that the ith row of MP (WP) gives the preferences of m_i (w_i). For example, $MP[i, j] = k$ if m_i's jth preference is w_k.

Two other useful matrices in a problem instance are the ranking matrices, denoted MR and WR. An entry in the men's ranking matrix, $MR[i, j]$, gives the ranking (position of preference) of w_j by m_i. The preference and ranking matrices play the roles of inverses for each other because $MP[i, MR[i, j]] = j$ and $MR[i, MP[i, j]] = j$. Given MP (WP), it is possible to determine MR (WR) completely, and vice versa.

We will use the notations MP, MR, WP, WR only when the problem instance associated with these matrices can be clearly determined from context. When there is a possibility of ambiguity, we use the notations MP_S, MR_S, WP_S, WR_S where S denotes a specific instance of the stable marriage problem.

A matching $M = \{(m_1, w_{i_1}), (m_2, w_{i_2}), \ldots, (m_n, w_{i_n})\}$ is a *stable marriage* if there does not exist an unmatched man-woman pair (m_i, w_j) such that both prefer each other to their partners in M. At least one stable marriage exists for any given problem instance. In most problem instances, there exists more than one stable marriage. Moreover, there are problem instances of size n where the number of stable marriages are exponential in n [IL86] [Kn76].

Gale and Shapley first proved that a stable marriage exists for any problem instance and gave the following algorithm to find it.

Gale-Shapley algorithm [GS62].
Initially, all men belong to the set B of bachelors. A man is removed from B by being engaged to a woman but may subsequently return to B if his fiancée breaks the engagement.

The algorithm iterates the following steps until B is empty.

Step 1: Select an m_i and remove him from B.
Step 2: m_i asks for an engagement from w_j, the woman highest on his preference which he has not previously asked.
Step 3: The reply of w_j falls into one of three cases:
a) if w_j is not engaged, she accepts m_i's request for engagement,
b) else if w_j is currently engaged to a man she prefers to m_i, she rejects the request and returns m_i to B,
c) otherwise w_j breaks her current engagement, returns her fiancé to B, and accepts m_i's request for engagement.

When B is empty, the matching represented by the engagements is a stable marriage.

The stable marriage derived from the algorithm is independent of the selection process used to pick m_i in Step 1. In a variation developed by McVitie and Wilson [MW71], m_i is selected such that i is minimized over all remaining men in B. If the previous iteration ended in Step 3(b) or 3(c), the McVitie-Wilson selection criterion picks the man last returned to B by that step. Otherwise, no current member of B has yet participated in the algorithm, and the man currently holding the smallest subscript in B is selected.

The stable marriage derived from the Gale-Shapley algorithm is *male-optimal.* It has been shown that no man can receive a better match in any other stable marriage for the same problem instance [GS62]. By reversing the roles of men and women, the algorithm finds the *female-optimal* stable marriage.

Step 2 of the algorithm is implemented efficiently in $O(1)$ time by considering entries of *MP* in order. Step 3 also runs in $O(1)$ time by looking up *WR* and keeping track of the current engagement of each woman. The worst-case asymptotic time complexity of the entire algorithm is bounded by the number of iterations and has been shown to be $O(n^2)$ [IT78] [KN76].

A man-woman pair (m_i, w_j) is *stable* if it is contained in some stable marriage. Gusfield [GU87] gives an $O(n^2)$ algorithm that finds all stable pairs in a problem instance. It is easy to adapt Gusfield's algorithm for use in determining if an arbitrary pair is stable.

We conclude this section with a lemma proved by Irving and Leather [IL86].

Lemma 1. (Irving and Leather)
In a given instance of the stable marriage problem, if man m and woman w are partners in some stable marriage S then

a) There is no stable marriage S' in which both m and w have worse partners.
b) There is no stable marriage S' in which both m and w have better partners.

Model of Computation

One of the bottlenecks of stable marriage algorithms is the pre-processing needed for the men's and women's preferences. If it is necessary to input or set up the preference and ranking matrices, then $\Omega(n^2)$ operations are required even before the algorithm proper begins its execution.

It is interesting to investigate if the stable marriage problem can be solved without examining all the preferences. Knuth, in one of the twelve research problems posed in his treatise on stable marriage [KN76], asks if there is an algorithm that finds a stable marriage, such that the *core* of the algorithm (as oppose to setting up the preferences) runs in $o(n^2)$ (less than $\Theta(n^2)$) time. In a related problem, Gusfield [GU87] asks if it is possible to determine in $o(n^2)$ time if an arbitrary complete matching is stable. Our results answer both questions in the negative.

The model of computation used in our investigation of lower bounds excludes the time needed to set up the preference and ranking matrices. Instead, we assume that the matrices MP, WP, MR, WR are already available. The algorithm may query any entry in these matrices and receive an answer in $O(1)$ time.

Adversary Strategy

Adversary argument is a general approach for showing lower bounds. When applied to a problem X, it can be modeled as a game involving two persons, the algorithm and adversary, making alternate moves. During her moves, the algorithm asks questions to gain information about X. Her overall goal is to minimize the number of questions needed to obtain sufficient information for solving X.

The adversary is obliged to answer the algorithm's question made during her previous move. His goal is to slow the algorithm by behaving uncooperatively, giving redundant answers whenever feasible. Although the answers must be incrementally consistent, he is free to construct a 'tricky' instance of X as the game proceeds. Benefiting from the assembled knowledge of questions already committed by the algorithm, he chooses replies that make the remaining task of the algorithm as time consuming as possible.

For every size n, our adversary uses a special instance of the stable marriage problem which we call the *canonical instance* and denote by C. The pair (m_n, w_n) is stable in C. There exists a large family of other problem instances that differ only slightly from C; yet, the differences are sufficient to cause (m_n, w_n) to be not stable. We show later how to construct such a problem instance which we called a *minimally non-canonical instance* and denote by $\sim C$.

The adversary's strategy is to answer queries according to the preference matrices of C. To conclude that (m_n, w_n) is stable, an algorithm must determine that no potential candidate for $\sim C$ agrees with those preferences in C that the adversary has revealed in answers to queries. However, C is constructed to provide a large number of possible $\sim C$s each derivable with only minor changes to C. Hence, the algorithm must make a large number of queries to eliminate all potential $\sim C$s, supporting our lower bound claim.

We now define the women's preference matrix WP_C. Entries in WP_C are defined by the function $WP_C[i, j] = j$ as illustrated in Figure 1.

$$\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & \dots & n \end{pmatrix}$$

Figure 1. Women's preference matrix WP_C.

We note a property of WP_C due to its construction. If a woman receives proposals from two different men, her decision always favors the man with the smaller subscript. This observation permits a simpler application of the Gale-Shapley algorithm that uses the McVitie-Wilson selection criterion. When m_i is selected in Step 1 for the first time, all men with subscripts greater than i have yet to participate in the algorithm. Hence, Step 3(c) never occurs because it requires that w_j be first engaged to m_k before receiving a proposal from m_i where $k > i$. In other words, a woman never changes her mind once she accepts an engagement. It follows that Steps 2 and 3 are reduced to scanning, in decreasing order, m_i's preferences and matching him with the first woman that is not engaged. We refer to this special version as the *restricted Gale-Shapley algorithm* in subsequent discussions.

The following lemma is another consequence of WP_C's construction.

Lemma 2.
Regardless of the men's preferences, any problem instance that has WP_C as the women's preference matrix yields exactly one stable marriage.

Proof: Any stable marriage is represented in WP_C by exactly one entry in each column. In particular, this is true of the male-optimal stable marriage S derived by applying the Gale-Shapley algorithm. Since S is male-optimal, if there exists another stable marriage S', every man that receives a different match now has a less preferable partner. By Lemma 1 (part a), every woman in S' has the same or a more preferable partner than in S. Therefore, the subscript of each woman's partner decreases or stays the same. However, this requires that some column in WP_C be represented in S' by more than one entry. We conclude that S' does not exist. ∎

Entries in the men's preference matrix MP_C fall into three groups. The first group, underlined in Figure 2, includes the first row, last row and tridiagonal entries of the remaining rows. The first and last rows consist of the integers 1 to n in increasing order. The tridiagonal entries of row i are the integers i, n, and $i-1$ in that order.

$$
\begin{array}{c}
\\ 1 \\ 2 \\ 3 \\ \vdots \\ i \\ \vdots \\ n-2 \\ n-1 \\ n
\end{array}
\begin{array}{c}
\begin{array}{ccccccccccccccc}
1 & 2 & 3 & 4 & 5 & \dots & i-1 & i & i+1 & i+2 & \dots & n-3 & n-2 & n-1 & n
\end{array}
\\
\left(\begin{array}{ccccccccccccccc}
\underline{1} & \underline{2} & \underline{3} & \underline{4} & \underline{\dots} & & & & & & & & \underline{n-2} & \underline{n-1} & \underline{n} \\
\underline{2} & \underline{n} & \underline{1} & 3 & 4 & \dots & & & & & & & & & n-1 \\
1 & \underline{3} & \underline{n} & \underline{2} & 4 & \dots & & & & & & & & & n-1 \\
\vdots & \vdots & & & \ddots & & & & & & & & & & \vdots \\
1 & 2 & \dots & & & i-2 & \underline{i} & \underline{n} & \underline{i-1} & i+1 & \dots & & & & n-1 \\
\vdots & \vdots & & & & & & & & \ddots & & & & & \vdots \\
1 & 2 & \dots & & & & & & & & n-4 & \underline{n-2} & \underline{n} & \underline{n-3} & n-1 \\
1 & 2 & \dots & & & & & & & & & n-3 & \underline{n-1} & \underline{n} & \underline{n-2} \\
\underline{1} & \underline{2} & \underline{3} & \underline{4} & \underline{\dots} & & & & & & & & \underline{n-2} & \underline{n-1} & \underline{n}
\end{array}\right)
\end{array}
$$

Figure 2. Men's preference matrix MP_C.

Of the remaining entries, the group left of the tridiagonals consists of integers 1 to $i-2$ in increasing order. The group right of the tridiagonals consists of integers $i+1$ to $n-1$ also arranged in increasing order.

Lemma 3.
(m_n, w_n) is a stable pair in C.

Proof: Apply the restricted Gale-Shapley algorithm to C. Scanning the ith row of MP_C, note that all entries to the left of i have values less than i. These entries represent those women who have already accepted engagements in previous rows. Hence, m_i's stable partner is w_i and $\{(m_1, w_1), (m_2, w_2), \dots, (m_n, w_n)\}$ is a stable marriage in C. ∎

Obtaining Non-canonical Instances

Starting with C, we obtain a $\sim C$ by selecting a row i of MP_C such that $3 \leq i \leq n-2$ and exchanging two special entries, l and r, in that row. All entries left of the tridiagonal are candidates for l but only those right of the tridiagonal with *values that differ from i by odd numbers* are candidates for r. Note that l is equal to its column number, and r's column number is $r+1$.

To formalize the above construction, we define, for each i, two sets of integers

$$L_i = \{\, x \mid 1 \leq x \leq i-2 \,\}, \quad \text{and}$$
$$R_i = \{\, x \mid i+1 \leq x \leq n-1 \quad \text{and} \quad x \not\equiv i \pmod 2 \,\}.$$

Then, for any i, l, and r satisfying $3 \leq i \leq n-2$, $l \in L_i$, and $r \in R_i$; we define $MP_{\sim C}[i,l] = r$, $MP_{\sim C}[i,r+1] = l$, and all other entries of $MP_{\sim C}$ and $WP_{\sim C}$ are equal to their corresponding entries in MP_C and WP_C.

Lemma 4.
(m_n, w_n) is not a stable pair in $\sim C$.

Proof: Apply the restricted Gale-Shapley algorithm to $\sim C$. Figure 3 illustrates the stable marriage that results.

- m_k is matched with w_k for $1 \leq k \leq i-1$ because these rows are unchanged from C.
- m_i is matched with w_r. Note that $r \not\equiv i \pmod 2$, which guarantees that there is an even number of rows between row i and row r.
- For $i+1 \leq k \leq r-1$, m_k is matched with w_k if $k \not\equiv i \pmod 2$ and m_k is matched with w_{k-2} if $k \equiv i \pmod 2$. Note that m_{r-2} is matched with w_{r-2} and m_{r-1} is matched with w_{r-3}.

The above discussion shows that $w_1, w_2, \ldots, w_{i-1}$ have accepted engagements in rows 1 to $i-1$; $w_i, w_{i+1}, \ldots, w_{r-2}$ have accepted engagements in rows $i+1$ to $r-1$ and w_r has accepted an engagement in row i. The subscripts of these women account for every entry left of the diagonal entry n in row r. Hence, m_r is engaged to w_n.

$\sim C$ has only one stable marriage by Lemma 2. Since w_n is married to m_r and not m_n in this marriage, (m_n, w_n) is not a stable pair. ∎

Under the rules of the adversary argument, the adversary cannot renege on a reply to an earlier query. Hence, the l and r values, which are exchanged during the construction of $\sim C$, cannot originate from entries that the algorithm had queried. However, the large number of valid choices of i, l and r gives us the following bound.

$$
\begin{array}{cc}
 & \begin{array}{ccccccccccccccccc}
1 & 2 & 3 & \dots & l & \dots & i-2 & i-1 & i & i+1 & i+2 & \dots & r-3 & r-2 & r-1 & r & \dots
\end{array} \\
\begin{array}{c}
1 \\ 2 \\ 3 \\ \vdots \\ i-1 \\ i \\ i+1 \\ i+2 \\ i+3 \\ i+4 \\ \vdots \\ r-2 \\ r-1 \\ r \\ \vdots
\end{array} &
\begin{pmatrix}
\boxed{1} & 2 & 3 & \dots & & & & & & & & & & & & & \\
\boxed{2} & n & 1 & \dots & & & & & & & & & & & & & \\
1 & \boxed{3} & n & \dots & & & & & & & & & & & & & \\
\vdots & & & \ddots & & & & & & & & & & & & & \\
1 & 2 & \dots & & & \dots & \boxed{i-1} & n & i-2 & \dots & & & & & & & \\
1 & 2 & \dots & & \boxed{r} & & \dots & i & n & i-1 & \dots & & & & & & \\
1 & 2 & \dots & & & & \dots & i-1 & \boxed{i+1} & n & i & \dots & & & & & \\
1 & 2 & \dots & & & & \dots & i-1 & \boxed{i} & i+2 & n & i+1 & \dots & & & & \\
1 & 2 & \dots & & & & & & \dots & i+1 & \boxed{i+3} & n & \dots & & & & \\
1 & 2 & \dots & & & & & & \dots & i+1 & \boxed{i+2} & i+4 & \dots & & & & \\
\vdots & \vdots & & & & & & & & & \vdots & \ddots & & & & & \\
1 & 2 & \dots & & & & & & & & & \dots & \boxed{r-2} & n & r-3 & \dots & \\
1 & 2 & \dots & & & & & & & & & \dots & \boxed{r-3} & r-1 & n & r-2 & \dots \\
1 & 2 & \dots & & & & & & & & & \dots & r-3 & r-2 & r & \boxed{n} & \dots \\
\vdots & \vdots & & & & & & & & & & & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\end{array}
$$

Figure 3. Stable marriage in $\sim C$.

Lemma 5.
If $n = 3k + 4$ for some integer $k \geq 1$, the minimum number of queries needed to eliminate all valid constructions of $\sim C$s is $\frac{3}{2}k(k+1)$.

Proof: To eliminate row i from participating in the construction of a $\sim C$, the algorithm must query either all of L_i or all of R_i. To eliminate all valid constructions of $\sim C$s, all rows must be eliminated.

$$\begin{array}{ccccccccccccc} |L_i| & = & i-2 & \leq & k & < & \lceil (2k+1)/2 \rceil & \leq & \lceil (n-i-1)/2 \rceil & = & |R_i| \\ & & & & & & \text{for } 3 \leq i \leq k+2, & & \text{and} & & \\ |L_i| & = & i-2 & \geq & k+1 & > & \lceil 2k/2 \rceil & \geq & \lceil (n-i-1)/2 \rceil & = & |R_i| \\ & & & & & & \text{for } k+3 \leq i \leq n-2. & & & & \end{array}$$

Therefore the minimum number of queries needed

$$\begin{aligned} &= \sum_{i=3}^{n-2} \min(|L_i|, |R_i|) \\ &= \sum_{i=3}^{k+2} |L_i| + \sum_{i=k+3}^{n-2} |R_i| \\ &= \sum_{i=3}^{k+2} (i-2) + \sum_{i=k+3}^{n-2} \lceil (n-i-1)/2 \rceil \\ &= \sum_{i=3}^{k+2} (i-2) + \sum_{i=k+3}^{3k+2} \lceil (3k+3-i)/2 \rceil \\ &= \sum_{j=1}^{k} j + \sum_{j=1}^{k} 2j \\ &= \frac{3}{2}k(k+1). \ \blacksquare \end{aligned}$$

Lower Bounds Results

We are now ready to state our main result.

Theorem.
Determining if an arbitrary pair is stable in a problem instance of size n requires $\Omega(n^2)$ time in the worst case.

Proof: Without loss of generality, we may assume that $n = 3k + 4$ for some integer $k \geq 1$; otherwise, we extend the problem instance by adding the appropriate number of men and women.

By Lemmas 3 and 4, we show that it is necessary to distinguish between C and $\sim C$ in order to determine if (m_n, w_n) is stable. By Lemma 5, any algorithm that distinguishes between C and $\sim C$ must make at least $\frac{3}{2}k(k+1) = \frac{3}{2}(\frac{n-4}{3})(\frac{n-4}{3}+1)$ queries. Hence, the number of queries necessary is $\Omega(n^2)$. ∎

Corollary 1.
The asymptotic time complexity for determining if an arbitrary pair is stable in a problem instance of size n is $\Theta(n^2)$.

Proof: Corollary 1 follows from the theorem and our earlier note that Gusfield's algorithm solves this problem in $O(n^2)$ time. ∎

Corollary 2.
The asymptotic time complexity for finding a stable marriage in a given problem instance of size n is $\Theta(n^2)$.

Proof: We noted earlier that Gale-Shapley algorithm runs in $O(n^2)$ time. The only stable marriage in C is different from the only stable marriage in $\sim C$ and $\Omega(n^2)$ queries are required to distinguish between them. ∎

Corollary 3.
The asymptotic time complexity for determining if an arbitrary complete matching $\{(m_1, w_{i_1}), (m_2, w_{i_2}), \ldots, (m_n, w_{i_n})\}$ is a stable marriage in a given problem instance of size n is $\Theta(n^2)$.

Proof: There is an obvious $O(n^2)$ algorithm for solving this problem described by Gusfield [GU87, p 127]. The matching $\{(m_1, w_1), (m_2, w_2), \ldots, (m_n, w_n)\}$ is a stable marriage in C but not in $\sim C$ and $\Omega(n^2)$ queries are required to distinguish between them. ∎

Conclusions

We have shown that the lower bound of $\Omega(n^2)$ holds for three stable marriage related problems. This lower bound is fundamental to stable marriage and holds for other related problems, such as, the stable 'roommates' problem [IR85] and an entire class of stable marriage problem variants defined as follows. Given an instance of the stable marriage problem X, we define a real-valued function V whose domain is the set of stable marriages in X. The problem of of finding a stable marriage M that maximizes (or minimizes) $V(M)$ has a lower bound of $\Omega(n^2)$, by an argument similar to that of Corollary 2. By varying the definition of V, we can formulate different stable marriage problem variants.

Most of the problems that we have described can be solved using existing algorithms that run in $O(n^2)$ time. Our lower bound implies that these algorithms are asymptotically optimal. However, a notable exception is the 'optimal' stable marriage problem [ILG87], which is a member of the class of problems we have just described. In this problem, two functions $um(i,j)$ and $uw(i,j)$ are defined for every pair (m_i, w_j). These functions represent respectively the 'unhappiness' ratings given by m_i and w_j to any stable marriage that pairs them together. The problem is to find a stable marriage M that minimizes $V(M)$, which is defined to be the sum of um's and uw's over all pairs in M. A solution represents an 'optimal' stable marriage in an egalitarian sense, which maximizes the overall happiness of all participants. An algorithm that runs in $O(n^4 \log n)$ time is given in [ILG87]; the interesting open question of whether there exists an asymptotically faster algorithm is also raised. Our results only show a lower bound of $\Omega(n^2)$.

References

[GS62] D. Gale and L. Shapley, *College admissions and the stability of marriage*, Amer. Math. Monthly, 69(1962), pp. 9–15.

[Gu87] D. Gusfield, *Three fast algorithms for four problems in stable marriage*, SIAM J. Comput., 16(1987), pp. 111–128.

[IL86] R. W. Irving and P. Leather, *The complexity of counting stable marriages*, SIAM J. Comput., 15(1986), pp. 655–667.

[ILG87] R. W. Irving, P. Leather and D. Gusfield, *An efficient algorithm for the "optimal" stable marriage*, J. Assoc. Comput. Mach., 34(1987), pp. 532–543.

[Ir85] R. W. Irving, *An efficient algorithm for the "stable roommates" problem*, J. Algorithms, 6(1985), pp. 577–595.

[It78] S. Y. Itoga, *The upper bound for the stable marriage problem*, Oper. Res., 29(1978), pp. 811–814.

[Kn76] D. E. Knuth, *Mariages Stables*, Les Presses de L'Université de Montréal, Montréal, 1976.

[MW71] D. G. McVitie and L. B. Wilson, *The stable marriage problem*, Comm. ACM, 14(1971), pp. 486–492.

[Wi72] L. B. Wilson, *An analysis of the stable marriage assignment problem*, BIT, 12(1972), pp. 569–575.