Dartmouth College
Dartmouth Digital Commons

Computer Science Technical Reports | Computer Science

12-1-2005

How hard is it to cheat in the Gale-Shapley Stable Matching Algorithm

Chien-Chung Huang
Dartmouth College

Follow this and additional works at: https://digitalcommons.dartmouth.edu/cs_tr

Part of the Computer Sciences Commons

Dartmouth Digital Commons Citation
Huang, Chien-Chung, "How hard is it to cheat in the Gale-Shapley Stable Matching Algorithm" (2005). Computer Science Technical Report TR2005-565. https://digitalcommons.dartmouth.edu/cs_tr/284

This Technical Report is brought to you for free and open access by the Computer Science at Dartmouth Digital Commons. It has been accepted for inclusion in Computer Science Technical Reports by an authorized administrator of Dartmouth Digital Commons. For more information, please contact dartmouthdigitalcommons@groups.dartmouth.edu.

How Hard is it to Cheat in the Gale-Shapley Stable Matching Algorithm?

Chien-Chung Huang
Technical Report TR2005-565
Department of Computer Science
Dartmouth College
villars@cs.dartmouth.edu

Abstract

We study strategy issues surrounding the stable marriage problem. Under the Gale-Shapley algorithm (with men proposing), a classical theorem says that it is impossible for every liar to get a better partner. We try to challenge this theorem. First, observing a loophole in the statement of the theorem, we devise a coalition strategy in which a non-empty subset of the liars gets a better partner and no man is worse off than before. This strategy is restricted in that not everyone has the incentive to cheat. We attack the classical theorem further by means of randomization. However, this theorem shows surprising robustness: it is impossible that every liar has the chance to improve while no one gets hurt. Hence, this impossibility result indicates that it is always hard to induce some people to falsify their lists. Finally, to overcome the problem of lacking motivation, we exhibit another randomized lying strategy in which every liar can expect to get a better partner, though with a chance of getting a worse one.

1 Introduction

Suppose that n men and n women seek a life-long partner. Each of them has a preference list of the members of the other sex and submits it to a centralized authority. In the spirit of making all the participants maintain a long-term relationship, the authority has to make sure that the matching does not contain a *blocking pair*: a blocking pair is a couple each of whom prefers some one else over the other. A matching is *stable* if there is no blocking pair in it. The goal of the authority, given the men's and women's preference lists, is to find a stable matching.

The above situation is the classical stable marriage problem first formulated by Gale and Shapley [4]. They not only showed that there always exists at least one stable matching, but also gave a time-optimal algorithm to find it. Henceforward, the stable marriage problem and its variants have never ceased to interest researchers in various areas. Some researchers are especially intrigued by strategies surrounding the stable marriage problem. Suppose the match-making mechanism is known beforehand, and all men's and women's preference lists are made public. Can a group of persons (of either sex) somehow falsify their lists to get better partners?

For the Gale-Shapley men-optimal algorithm, some studies have partly answered the question. If women are allowed to submit incomplete lists (i.e., they can declare some men unacceptable), they can force a men-optimal matching into a women-optimal one [5]. For men, researchers reached the opposite conclusion: honesty is the best policy [3, 11]. We state the theorem formally, as it inspires this work and is the key to the proof of our several results.

Theorem 1 *A subset of men cannot falsify their preference lists so that every one of them gets a better partner than in the Gale-Shapley men-optimal algorithm.*

This theorem is significant in its generality and implications.[1] First, it indicates that when we adopt the Gale-Shapley men-optimal algorithm, it is futile for a man to lie when all other people are honest, since the most he can hope for is simply the same partner that he got by being honest. Second, collectively, men need not bother about proposing using different matching algorithms in the hope that they are more easily maneuverable by lying. Whatever algorithms we adopt, there does not exist a stable matching (given their falsified lists) in which all liars get better partners. Hence, the men's ideal (and simple) plan should be to stick to the Gale-Shapley algorithm and be honest. Using the game-theoretic terminology, Theorem 1 tells us that the Gale-Shapley algorithm intrinsically ensures that honesty is the dominant strategy.

This work tries to challenge this seemingly unassailable theorem. Is it really impossible to contrive some strategies to get around it? Indeed, a first hint emerges when we observe the wording of the theorem more carefully. It is impossible that *every one of them gets a better partner*. But this statement does not rule out the possibility that *some subset of liars gets a better partner while the others get the same partners*.[2] As we are going to show, it is possible to form a coalition strategy following this observation. Moreover, we prove this is the *only* possible strategy to cheat in which none of the liars is worse off.

Our second attempt tries to broaden the breach by randomizing the coalition strategy. We surmise that a randomized strategy should ensure that everyone has the chance to improve his partner and no one is going to be worse off. Such a randomized strategy would encourage men to falsify their preference lists. However, we reach an impossibility result saying that such a randomized strategy is unrealizable. Therefore, paradoxically, our coalition strategy strengthens Theorem 1 inadvertently.

Finally, we relax the requirement of the randomized strategy. Liars are allowed to be worse off sometimes. We then present another randomized strategy in which every liar can *expect* to get a better partner. Thus, in an amortized sense, our third assault on Theorem 1 did "succeed".

The main contribution of this work is the re-examination of the classical theorem and its associated strategy issues. To our knowledge, ours is the first attempt in proposing men-lying strategies (both deterministic and randomized).

The outline of this paper is as follows. In Section 2, we observe the interaction between the preference lists and the matching. Section 3 formally presents the coalition strategy. In Section 4, we delve deeper into the the motivational issues of the coalition strategy, showing that there always exist some men who do not gain by lying. In Section 5, we exhibit another randomized lying strategy in which men on the average can get a better partner. Section 6 concludes and discusses some related work.

2 Falsifying Preference Lists

In this section, we observe the interaction between falsified lists and the resulting matchings. Our primary finding is that to get a matching in which every man is *at least* as well off as if everyone had been honest, the liars must lower the ranks of some highly-ranked women in their lists.

Before plunging into the technical details, we first establish some notation and terminology and give some background. Throughout this work, we assume that the Gale-Shapley men-optimal algorithm is used and we know all preference lists. We denote the sets of men and women by $\mathcal{M}$ and $\mathcal{W}$, both of size n. When everyone is honest, M_0 and M_z are the men-optimal and women-optimal matchings; M_s means the matching when some subset of people lie. For any matching M and any subset of people $S \subseteq \mathcal{M} \bigcup \mathcal{W}$, the collection of partners of people in S is $M(S)$. For example, $M_0(m)$ is the partner of man m in the men-optimal matching. We express the fact that man m prefers woman w over woman w' by $w \succ_m w'$.

[1] In fact, a stronger theorem states that even if a coalition of men and women lie together, they cannot all expect all to get better partners [2].

[2] This fact is also pointed out in [7].

Every man and woman has a strictly ordered preference list of size n. For man m, we define two sets $P_L(m) = \{w | w \succ_m M_0(m)\}$ and $P_R(m) = \{w | M_0(m) \succ_m w\}$. More colloquially, we say the women in $P_L(m)$(or $P_R(m)$) are on the left (right) of man m's list. A matching M "at least as good as" another matching M' is denoted as $M \succeq M'$, so that for every man m, $M(m) \succeq_m M'(m)$. If, besides $M \succeq M'$, there exists at least one man m such that $M(m) \succ_m M'(m)$, we write $M \succ M'$.

Finally, suppose A is a set of distinct objects, then $\pi(A)$ is the set of all $|A|!$ permutations. $\pi_r(A)$ means any one of the $|A|!$ permutations.

Our first Lemma hints at the necessary ingredient in men's falsified lists to get a better outcome for men. It motivates our coalition lying strategy, as will be formalized in the next section.

Lemma 2 *For a subset of men $S \subseteq \mathcal{M}$, if every member $m \in S$ submits the falsified list $(\pi_r(P_L(m) - X), M_0(m), \pi_r(P_R(m) \cup X))$, $X \subseteq P_L(m)$, then $M_s \succeq M_0$.*

Proof: We proceed by contradiction. In M_s, suppose some man m gets a worse partner than $M_0(m)$. Without loss of generality, assume that during the execution of the algorithm with true lists, m is the first person rejected by his M_0-partner. At that moment, $M_0(m)$ must be engaged to some other man m' who is higher up in her preference list. Since m' has not yet been accepted by his M_0 partner, he must prefer $M_0(m)$ over $M_0(m')$. Therefore, $(m', M_0(m'))$ must be a blocking pair in M_0. ∎

Lemma 2 implies that men can safely shift some women from the left to the right of their preference lists. In the worst case, the outcome is still the same as M_0. The question is whether this kind of shifting can lead to better results for men. In the next section we shall answer in the affirmative. Interestingly, Lemma 2 also has an intuitive interpretation: if some men know beforehand that they have no chance of getting some women, they may as well avoid proposing to them. Doing this, they do not run any risk of getting a worse partner and may help others get a better one.

It is natural to ask the analogous question of Lemma 2: how about shifting some women from the right to the left of men's preference lists? Intuitively, it is a dangerous move, because men will now first propose to some women they do not really like. In general, shifting women from the right to the left is more unpredictable in the outcome, but sometimes useful strategies follow this idea. We discuss one possible strategy in Section 5 and some others in the Appendix. For our purpose at this moment, we only show that it is impossible that by shifting women from the right to the left men's list, every man can be at least as well off as when they are honest.

Lemma 3 *For a subset of men $S \subseteq \mathcal{M}$, if every member $m \in S$ submits the falsified list $(\pi_r(P_L(m) \cup X), M_0(m), \pi_r(P_R(m) - X))$, $X \subseteq P_R(m)$, it is impossible that $M_s \succ M_0$.*

Proof: We again seek a contradiction. Suppose in M_s, every man is at least as well off as in M_0 and at least one man m strictly prefers $M_s(m)$ over $M_0(m)$. For every man who falsifies his list, we make him submit the original true list, then in M_0, at least one man is worse off, which contradicts Lemma 2. ∎

We next give another lemma, which goes a long way toward explaining why Lemma 2 is a useful lying stratagem.

Lemma 4 *For a subset of men $S \subseteq \mathcal{M}$, if every member $m \in S$ submits the falsified list $(\pi_r(P_L(m)), M_0(m), \pi_r(P_R(m)))$, then $M_s = M_0$.*

Proof: We can use the same argument in the proof of Lemma 2 to show that no man will ever be rejected by his M_0-partner. Hence, permuting the right portion of the men's preference lists will not cause men to be worse off. However, the permutation on the left portion of the preference lists might cause some men to be better off in M_s than in M_0. We have to eliminate this possibility.

M_0-matching		M_s-matching	
Men's List	Women's List	Men's (Falsified) List	Women's List
A: ab**e**dc	a: **C**BDAE	A:a**e**dcb	a: **C**BDAE
B: be**d**ac	b: **D**EABC	B:**b**edac	b: DEA**B**C
C: eb**a**cd	c: BCD**E**A	C:eb**a**cd	c: BCD**E**A
D: da**b**ce	d: A**B**CED	D:**d**abce	d: ABCE**D**
E: edb**c**a	e: **A**BECD	E: e**c**abd	e: **A**BECD

The Match Ranking for men in M_0: A(e,3), B(d,3), C(a,3), D(b,3), E(c,4)

The Match Ranking for men in M_s: A(e,3), B(b,1), C(a,3), D(d,1), E(c,4)

Figure 1: Men A and E falsify their lists to help men B and D get a better partner. Falsified lists are underlined.

Suppose there exists a nonempty subset $B \subset \mathcal{M}$ such that each man $m \in B$ is better off in M_s than in M_0. Given the falsified lists of men, the stability of M_s implies that every man $m \in B$ is preferred by his partner $M_s(m)$ over any other man $m' \in \mathcal{M} - B$ who puts $M_s(m)$ on the left of his preference list. In any execution of the Gale-Shapley algorithm with the true preference lists, the men in B must be rejected by their M_s-partners, and this rejection can be caused only by another man $m' \in B$. Moreover, after this rejection, his M_s-partner can be engaged only to men in B. Without loss of generality, assume that m is the last person in B who is rejected by his M_s-partner. At the point of this rejection, all the M_s-partners of men in B except $M_s(m)$ must have been engaged, and only to men in B. However, the rejection of m implies that $M_s(m)$ is also engaged to another man in B. Hence, $|B|$ women are engaged to $|B-1|$ men when the last rejection takes place, and we reach the desired contradiction. ∎

Lemma 4 tells us that it is vain to simply permute the left and/or the right portion of the men's preference lists. For something to really happen, the lists have to be altered more drastically. Therefore, from Lemma 2, 3, and 4, we know that for men to get better partners, it must involve some men shifting women from the left to the right of their lists.

3 Coalition Strategy

In this section, we first give an example to illustrate the intuition. We then formally present the coalition strategy.

An Example

Consider the example shown in Figure 1. We make two observations here. First, a man cannot improve his partner by lying alone (as Theorem 1 suggests). He has to have some "collaborators" and exchange their partners. If man B wants to be matched to woman b, one possibility is that he and man D exchange partners and become happier together. On the other hand, in this example, it is impossible for man E to improve, because his M_0-partner, woman c, is not on the left of any other person's preference list.

Second, continuing the preceding example, for men B and D to be matched to women b and d respectively, men A, C, and E, when proposing all the way down to their M_0-partners, should avoid breaking the "balance" of men B and D and women b and d. For example, once man A proposes to woman b, he will cause man B to be rejected by woman b; on the other hand, if man C makes proposal to b, it does not matter, as man B is higher up than man C in woman b's list. As predicted by Theorem 1, the men falsifying their lists (men A and E) do not all get a better partner, but they do help other people (men B and D) get a better one.

Coalitions

We now formally explain the coalition strategy. A coalition is comprised of two parts: *cabal* and *accomplices*. The cabal is circularly ordered and each man in the cabal prefers his predecessor's M_0-partner to his own. To be more precise,

Definition 5 *The cabal of a coalition $K = (m_1, m_2, \cdots, m_{|K|})$ is a list of men such that each man m_i, $1 \leq i \leq |K|$, prefers $M_0(m_{i-1})$ to his own partner $M_0(m_i)$, indices taken module $|K|$.*

Having formed the cabal, the next step is to enlist the help of its *accomplices*. These accomplices have to shift the women preferred by the people in the cabal to the right of their lists (as implied by Lemma 2), if they are preferred by the women over the people in the cabal. The formal definition is as follows:

Definition 6 *The set of accomplices of cabal $K = (m_1, m_2, ...m_{|K|})$ is a set of men $A(K) \subseteq \mathcal{M}$ such that $m \in A(K)$ if*

1. $m \notin K, m_i \in K, M_0(m_i) \succ_m M_0(m), m \succ_{M_0(m_i)} m_{i+1}$, *or*
2. $m = m_j \in K, m_i \in K, i \neq j, M_0(m_i) \succ_{m_j} M_0(m_{j-1}), m_j \succ_{M_0(m_i)} m_{i+1}$.

Note that cabal K and its accomplices $A(K)$ might not be disjoint, i.e., the people in the cabal might have to falsify their lists as well. An immediate consequence of Theorem 1 is that $A(K) \cup K \neq K$.

We can now present the main result of this section.

Theorem 7 *Coalition Strategy: If in a coalition $C = (K, A(K))$, each accomplice $m \in A(K)$ submits the falsified list $(\pi_r(P_L(m) - X), M_0(m), \pi_r(P_R(m) \cup X))$, and if*

- $m \in A(K) - K, X = \{w | w = M_0(m_i) \in M_0(K), m \succ_w m_{i+1}\}$
- $m = m_j \in A(K) \cap K, X = \{w | w = M_0(m_i) \in M_0(K), w \succ_{m_j} M_0(m_{j-1}), m_j \succ_w m_{i+1}\}$,

then in the resulting M_s-matching, $M_s(m) \succ_m M_0(m)$ for $m \in K$ and $M_s(m) = M_0(m)$ for $m \notin K$.

Proof: As in Lemma 2, no man is going to be rejected by his M_0-partner, since men only shift some women from the left to the right of their lists. Moreover, no man m_i in K is going to be rejected by his preferred partner $M_0(m_{i-1})$, since all the accomplices have altered their lists. Finally, men not in the cabal can get only their M_0-partners and men in the cabal can get only their preferred partners. As in the proof of Lemma 4, after assuming there exists a subset of men in M_s who get even better partners, we can use a pigeonhole argument to refute its existence. ∎

The coalition strategy works, and moreover, it is the *only* strategy that has the nice property of ensuring some men are better off and every liar is at least as well off as before.[3]

Theorem 8 *The coalition strategy is the only way for men to falsify their lists such that in the resulting matching some men are better off and every liar is at least as well off as when he is honest.*

[3]Throughout the work, we use the term "coalition strategy" slightly loosely. When we say the "coalition strategy", we mean that the accomplices have shifted the necessary women from the left to the right of their lists to help men in the cabal. Indeed, their lists can allow more changes while still ensuring the men in the cabal get their desired partners as discussed in the Theorem 7. We discuss the details in the Appendix

Proof: We proceed by contradiction. Suppose there exists another strategy for men such that some men can be better off at the expense of other honest men and all liars are at least as well off as when they are honest. If some man m (whether he is honest or not) is better off by being matched to the partner of some honest man m', i.e. $M_s(m) = M_0(m')$, while the honest man m' is worse off. We claim that $(m', M_0(m'))$ must be a blocking pair in M_s, because (1) the stability of M_0 implies that $m' \succ_{M_0(m')} m$, and (2) since m' is honest, $M_0(m') \succ_{m'} M_s(m')$. ■

Theorem 8 has an important implication: Liars, if intending to help other men (or themselves) get a better partner, either have to adopt the coalition strategy (in which no one gets hurt) as defined in Theorem 7, or must accept a worse partner for themselves. This observation prompts us to devise another strategy in Section 5.[4]

The algorithms for finding the coalitions (cabals) can be found in the Appendix. We discuss some other theoretical implications that directly follow from the coalition strategy.

Cabalists and Hopeless Men

Based on the preference lists and the M_0-matching, a large number (which can be exponential) of coalitions may exist. We define man m to be one of the *cabalists* $\mathcal{K}$ if he belongs to any one of the cabals of the coalitions; otherwise, he is one of the *hopeless men* $\mathcal{H}$. By this definition, men fall into two categories: $\mathcal{M} = \mathcal{K} \bigcup \mathcal{H}$ and $\mathcal{K} \bigcap \mathcal{H} = \emptyset$. Apparently, hopeless men cannot benefit from utilizing our coalition strategy. The following lemma is important in proving our next major result.

Lemma 9 *Whatever the true preference lists, there always exists at least one hopeless man, i.e.,* $\mathcal{H} \neq \phi$.

Proof: We first claim that if woman w is the last woman receiving a proposal, then (1) she has not received any other proposal before, and (2) she is not in the left portion of any man's preference list. If this is not so, then when the last proposal is made to w, she will either reject the proposer or dump her former partner. In both cases, this "last" proposal will not terminate the algorithm.

Since the last woman w receiving a proposal is not on the left of any man's preference list, $M_0(w)$ cannot belong to any cabal. Hence he must be one of the hopeless men. ■

Lemma 9 tells us that at least one man does not have the incentive to lie. A little by-product of Lemma 9 is an easy proof of weak pareto-optimality of M_0, which has been shown before [6, 11].

Corollary 10 *There does not exist a matching* M^**, stable or not, such that every man gets a strictly better partner than in* M_0.

Proof: Since the last woman w receiving the proposal is not on the left of any man's preference list, there cannot be a matching in which every man has a better partner and one of them is matched to w. ■

4 Impossibility of Forming Leagues

Our coalition strategy has one unsatisfactory aspect: Theorem 1 ordains that for every coalition, there is at least one accomplice who does not gain from lying and hence has little motivation of doing so. It is tempting to push the frontier further. Can we devise some stratagem such that everyone has incentive to perjure himself? In this section, we show that even with a randomized strategy, we still cannot overcome the problem of lacking motivation.

We formulate what is a successful randomized strategy for men.

[4] However, Theorem 8 *does not* imply that honest men will never be worse off when other people are not playing fair. It is possible that they get a worse partner. We discuss this issue in the Appendix.

Definition 11 *Suppose we have a league of men $L \subseteq \mathcal{M}$. Each man $m_i \in L$ has a set of possible preference lists $s_i = \pi(|\mathcal{W}|)$. We define a joint probability distribution $F : s_1 \times s_2 \cdots \times s_{|L|} \rightarrow [0, 1]$. A randomized strategy is successful for the league if for every man $m_i \in L$:*

- *(Positive Expectation): $E[Rank(M_s(m_i))] > Rank(M_0(m_i))$.*
- *(Elimination of Risk): If in event E, $Rank(M_0(m_i)) > Rank(M_s(m_i))$, then $P(E) = 0$.*

The first requirement is obvious, since we aim to encourage people to falsify their lists. We consider the second requirement as a desirable feature of any randomized strategy. Ideally, we should ensure that none of the liars is going to be worse off. In real world, the second requirement translates into the fact that rational men are risk-averse.

Based on Theorem 8, the two requirements imply that the only choice is to employ the coalition strategy. We can randomly pick some coalition contained in the league and realize the strategy accordingly. The problem then boils down to whether we can find a union of coalitions $C_i = (K_i, A(K_i))$ such that $L = \bigcup_i K_i = \bigcup_i A(K_i)$. In other words, in this league, all accomplices belong to the cabal of some coalition and have the chance to improve his partner (and hence the incentive to lie).

To a certain extent, a league circumvents Theorem 1 and creates the possibility that every liar can improve his partner (in a randomized sense). However, as we are going to show, a league cannot exist.

Theorem 12 *For any coalition $C = (K, A(K))$, at least one of its accomplices is a hopeless man, i.e., $A(K) \bigcap \mathcal{H} \neq \emptyset$.*

Proof: We first consider *maximal* coalitions and then go on to more general cases. A coalition $C = (K, A(K))$ is maximal if $K = \mathcal{M} - \mathcal{H}$. For every man m_i in the cabal of this maximal coalition, we move his preferred partner $M_0(m_{i-1})$ in the cabal to the front of his list and his M_0-partner $M_0(m_i)$ to the second place. Note that due to Lemma 2, after this alternation of the lists, a man in the cabal can be matched only to either his original M_0-partner or his preferred partner in the cabal.

Arrange the proposal sequence of the Gale-Shapley algorithm in the following way: all men in $\mathcal{M} - \mathcal{H}$ propose first and are temporarily engaged to their preferred partners in the cabal. In the resulting matching, Theorem 1 tells us that it is impossible that every liar gets a better partner, so at least one person m_j in the cabal is matched to his M_0-partner $M_0(m_j)$; consequently, m_{j+1} also can be matched only to his original M_0-partner $M_0(m_{j+1})$ and so forth. The only way to break the "balance" of this cabal is that some hopeless man m^* (there exists at least one hopeless man, as indicated by Lemma 9) proposes to some woman in the coalition and he is preferred by this woman over the men in the cabal. Hence, m^* must be one of the accomplices in this coalition.

If the coalition C is not maximal, i.e., $|K| < |\mathcal{M} - \mathcal{H}|$, we still can apply the above argument, with a little more complication. First, choose some cabalist m not in K, and move his M_0-partner to the front of his preference list. Then, for all other cabalists m_k, if $M_0(m) \succ_{m_k} M_0(m_k)$, shift her to the right of his list. We claim that after this step, m becomes hopeless and the resulting M_s-matching would still be the same as M_0. The reasons are as follows: (1) If there exists any other cabal K' one of whose members prefers $M_0(m)$, then the coalition containing the cabal K' cannot succeed, because m will not be rejected by his M_0-partner, who is on the front of his list. (2) Cabals other than K not involving m also cannot succeed, because all we have done is to shift $M_0(m)$ to the right of other men's preference lists. If such a coalition is to succeed, the accomplices of the coalition have to shift the preferred women in the cabal of the coalition to the right of their lists. But $M_0(m)$ is not one of them.

By applying the above argument repeatedly, we can make all cabalists in $\mathcal{M} - \mathcal{H} - K$ become hopeless. For the men in the cabal K (which is now a maximal coalition), use the same argument we have used before: for each m_i, shift $M_0(m_{i-1})$ and $M_0(m_i)$ to the first two places in his list. Let all men in K propose first. The "balance" of K can

be broken only by some true hopeless men (those originally in $\mathcal{H}$, instead of those false ones we created, because the latter will only propose to his M_0-partner and stop). By the above argument, we get the conclusion that every coalition has at least one accomplice who is a hopeless man. ∎

By Theorem 12, we know that an all-win league is impossible. A hopeless man never improves by the coalition strategy, which means that he can never attain the first requirement in Definition 11. Combining Theorem 8 and 12, we derive our major result in this section:

Theorem 13 *It is impossible to find a league, and a successful randomized strategy as defined in Definition 11 cannot be realized.*

A by-product of Theorem 12 is the following corollary.

Corollary 14 *Remove some subset S of men and their corresponding partners $M_0(S)$. If we apply the Gale-Shapley algorithm to $(\mathcal{M} - S, \mathcal{W} - M_0(S))$, then the resulting match would be (1) at least as good as the M_0-matching, except for those removed pairs, if $S \subset \mathcal{M}$; and (2) identical with the original M_0 matching, except for those removed pairs, if $S \subset \mathcal{M} - \mathcal{H}$.*

5 In Pursuit of Motivation

In this section, we show it is possible to devise a randomized strategy in which a group of liars can *expect* to get a better partner. The crucial point is that these liars must be willing to take the risk of getting a worse partner. Recalling the randomized strategy definition in the preceding section, we find the second requirement for a successful strategy a little too stringent: everyone participating in the cheating can *never* be worse off. Relaxing this requirement allows us to achieve the goal of motivating men to cheat.

We first introduce another lying strategy.

Lemma 15 *Victim Strategy: Suppose $M_0(m) \succ_m M_0(m')$, $M_0(m) \succ_{m'} M_0(m')$, and for all m_i, $M_0(m') \in P_L(m_i)$ implies $m \succ_{M_0(m')} m_i$. If m submits the falsified list $(\pi(P_L(m) \cup M_0(m')), M_0(m), \pi(P_R(m) - M_0(m')))$, then in the resulting matching M_s:*

1. *For m (the victim), $M_s(m) = M_0(m')$;*
2. *For m' (the benefiter), $M_s(m') \succ_{m'} M_0(m')$;*
3. *For men $m_i \neq m, m'$, $M_s(m_i) \succeq_{m_i} M_0(m_i)$.*

Proof: We construct a stable matching M^* as follows: Retain all the couples in M_0 except exchange the partners of m and m'.

We claim that the constructed M^* is stable, since every man, except m, has either the same or a better partner. For m, he also gets a "better" partner, since $M_0(m)$ is now on the left of his perjured preference list. And there is no danger of the existence of a blocking pair containing $M_0(m')$, since m is more favored by $M_0(m')$ than any other men putting her on the left of his list.

If the constructed M^* is not men-optimal, then the true M_s will still have the stated properties. Men-optimality of M_s ensures that every man gets the best possible partner among all stable matchings. The only exception is m, who can not get a better partner than $M_0(m')$ even in the true men-optimal matching, because of Theorem 1. ∎

M_0-matching		M_s-matching	
Men's List	Women's List	Men's (Falsified) List	Women's List
A: bd**a**ce	a: B**A**DCE	A:bd**c**ae	a: BAD**C**E
B: cd**b**ae	b: C**B**ADE	B:cd**b**ae	b: C**B**ADE
C: ad**c**be	c: A**C**BED	C:**a**dcbe	c: **A**CBED
D: a**e**bcd	d: **E**DBCA	D:a**e**bcd	d: **E**DBCA
E: **d**abce	e: **D**BECA	E: **d**abce	e: **D**BECA

The Match Ranking for men in M_0: A(a,3), B(b,3), C(c,3), D(e,2), E(d,1)
The Match Ranking for men in M_s: A(c,4), B(b,3), C(c,1), D(e,2), E(d,1)

Figure 2: An example: Man A shifts woman c from the right to the left of his list. He gets woman c and man C gets woman a. Note man B (C) also can use the same strategy to help man A (B).

A simple example for the victim strategy can be found in Figure 2. Unlike the coalition strategy which might require a large number of accomplices to cooperate, the victim strategy involves only one person (the victim), and under some circumstances, he can help a large number of men get a better partner. But the problem is the practicality of the victim strategy: where can we find people with such a self-sacrificing spirit? The randomness of the victim strategy makes possible that some men be willing to play the role of victim (occasionally).

Definition 16 *An alliance is a list of men $A = (m_1, m_2, ...m_{|A|})$, $1 \leq i \leq |A|$, indices taken modulo $|A|$, such that m_i can play the role of victim to help m_{i-1} get a better partner. In a pure strategy, some subset of the men in the alliance make this sacrifice. The alliance is considered successful if there is a probability distribution on pure strategies such that for every $m_i \in A$, $E[Rank(M_s(m_i))] > Rank(M_0(m_i))$.*

Depending on whether we wish to make the arrangement as fair as possible, or to maximize expected total partner rank, selecting the probability distribution can be sophisticated problem. Here we present an easy example. As shown in Figure 2, a successful alliance is composed of men A, B and C. Man A (or B, or C) can play the role of victim to help man C (or A, or B). Suppose we assign the probability of $1/3$ to each one of them to play the victim; then the expected rank of their partner would be $8/3$, which is an improvement.

6 Conclusion and Related Work

In this work, we propose a variety of lying strategies, both deterministic and randomized, for men in the Gale-Shapley men-optimal algorithm. We also strengthened the classical theorem stating that honesty is the best policy. Even with a randomized strategy, this theorem still holds. The only way to circumvent this theorem is that the liars must be willing to take some risk.

An important question related to the proposed strategies is how likely we can exploit them. Some simulation study or real-world case analysis would be helpful. A more significant result would be to do some probabilistic analysis, as Immorlica and Mahdian have done in [9].

An even more interesting issue is the situation that we only have partial information about preference lists. This setting increases the difficulty of framing strategies considerably. We leave this as an open question.

Related Work

The stable marriage problem, due to its theoretical appeal and practical applications, has spawned a large body of literature. For a summary, see [7, 10, 12]. Concerning the strategy aspect, several papers have shown the futility of men lying [3, 6, 11]. Immorlica and Mahdian [9] showed that if men have preference lists of constant size while

women have complete lists and both are drawn from an *arbitrary* distribution of permutation lists, then the chance of women benefiting from lying is vanishingly small. If women can submit truncated lists, Teo et al. [15] suggested a women-lying strategy, in which a woman can get her women-optimal partner when the men-optimal algorithm is run. They also show that if women are only allowed to permute their lists, then they might not get the women-optimal partner. We also have found a different women-lying strategy for the men-optimal algorithm. Our strategy only involves permuting women's lists (instead of relying on truncating them). Due to the space constraint, we refer the details to [8]. About permuting men's preference lists to manipulate the outcome of the matching, there is one small example in the book of Gusfield and Irving [7, P.65]. They attribute the example to Josh Benaloh.

References

[1] David Abraham, Katarína Cechlárová, David Manlove, and Kurt Mehlhorn. Pareto optimality in house allocation problems. In *Proceedings of the Fifteenth Annual International Symposium on Algorithms and Computation (ISSAC)*, pages 3–15, 2004.

[2] Gabrielle Demange, David Gale, and Marilda Sotomayor. A further note on the stable matching problem. *Discrete Applied Mathematics*, 16:217–222, 1987.

[3] Lester Dubins and David Freedman. Machiavelli and the Gale-Shapley algorithm. *American Mathematical Monthly*, 88:485–494, 1981.

[4] David Gale and Lloyd Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.

[5] David Gale and Marilda Sotomayor. Ms. Machiavelli and the stable matching problem. *American Mathematical Monthly*, 92:261–268, 1985.

[6] David Gale and Marilda Sotomayor. Some remarks on the stable matching problem. *Discrete Applied Mathematics*, 11:223–232, 1985.

[7] Dan Gusfield and Robert Irving. *The Stable Marriage Problem*. The MIT Press, 1989.

[8] Chien-Chung Huang. Cheating strategies in the Gale-Shapley algorithm. Unpublished manuscript.

[9] Nicole Immorlica and Mohammad Mahdian. Marriage, honesty, and stability. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 53–62, 2005.

[10] Donald Knuth. *Mariages stables et leurs relations avec d'autre problèmes combinatoires*. Les Presses de l'université de Montréal, 1976.

[11] Alvin Roth. The economics of matching: Stability and incentives. *Mathematics of Operations Research*, 7(4):617–628, 1982.

[12] Alvin Roth and Marilda Sotomayor. *Two-sided matching: A study in game-theorectic modeling and analysis*. Cambridge University Press, 1990.

[13] Lloyd Shapley and Martin Shubik. On cores and indivisibility. *Journal of Mathematical Economics*, 1:23–37, 1974.

[14] Robert Tarjan. Depth first search and linear graph algorithms. *SIAM Journal on Computing*, 1(2):146–160, 1972.

[15] Chung-Piaw Teo, Jay Sethuraman, and Wee-Peng Tan. Gale-Shapley stable marriage problem revisited: Strategic issues and applications. *Management Science*, 47:1252–1267, 2001.

A Shifting Women from the Right to the Left of Men's List

We investigate the results of shifting women from the right to the left of men's lists under the Gale-Shapley men-optimal algorithm. We have seen how the victim strategy uses this idea. The following discussion concerns two other issues: (1) Under what circumstances can we maintain the original M_0-matching? (2) How can malicious men exploit this idea to hurt other men?

A.1 Maintaining M_0-matching

Lemma 17 *Suppose man m submits the falsified list* $(\pi_r(P_L(m) \cup M_0(m')), M_0(m), \pi_r(P_R(m) - M_0(m')))$. *The resulting matching* $M_s = M_0$ *if:*

- $m' \succ_{M_0(m')} m$, *and*
- *No new possible coalition whose cabal involving m is formed.*

Proof: Using the same argument in Lemma 2, we can show that no man is going to be rejected by his M_0-partner. The only possible exception is m' due to the change that m has done to his list, but the statement of the lemma disallows that $M_0(m')$ prefers m over m'.

Consider the partner of m in M_s-matching. Due to Theorem 1, he can be matched only to $M_0(m')$ or $M_0(m)$. If the former case happens, then it must be some coalition whose cabal involving m is successfully achieved, but again this is impossible as the lemma disallows that some new coalition is formed.

Finally we argue that no other coalition whose coalition not involving m can be achieved in M_s. Assume the opposite. If some coalition C is achieved, then m shifting $M_0(m')$ again back to the right will cause M_s to become M_0, in which some men are worse off. This is in contradiction to Lemma 2. ■

Lemma 17 explains why in Note 3 we mentioned that we use the term "coalition strategy" loosely. Men are allowed to shift some women from the right to the left without changing the outcome as defined in Theorem 7.

A.2 Malicious Men Hurt Other Men

Lemma 18 *Suppose man m submits the falsified list* $(\pi_r(P_L(m) \cup M_0(m')), M_0(m), \pi_r(P_R(m) - M_0(m')))$. *If* $m \succ_{M_0(m')} m'$, *then* $M_0(m') \succ_{m'} M_s(m')$.

Proof: If man m really proposes to $M_0(m')$, then m' will be worse off, since m ranks higher than m' in $M_0(m')$'s list. Hence, if m' proposes to $M_0(m')$, he is bound to be rejected. The only other possibility is that m' never proposes to $M_0(m')$, ending up with someone even better for him. This implies that a coalition whose cabal including m' is successfully formed and there must be some other man m'' in the cabal who ends up with $M_0(m')$. But m'' is bound to be rejected in favor of m, since for $M_0(m')$, $m \succ_{M_0(m')} m' \succ_{M_0(m')} m''$, hence the coalition is sure to fail.

If m does not propose to $M_0(m')$, then it implies that he gets a better partner, which contradicts Theorem 1. ■

Note that Lemma 18 does not say anything about the result of m: he might have the same partner or be worse off. A more refined strategy ensures that a malicious man can hurt other men without getting a worse partner himself. And this is the reason in the Note 4 we state that being honest is not always safe for you. Details can be found in [8].

B Algorithms for the Coalition Strategy

In this section, we discuss some algorithmic questions arising out of our the coalition strategy. In particular, we are concerned about the following:

1. How to identify cabalists and hopeless men?
2. For any man, which women are possible for him by the coalition strategy?
3. What is the best possible lying strategy for men as a whole?

All three questions boil down to how to identify the cabal of the coalitions. Once the cabal is found, we can find its necessary accomplices in linear time by scanning through men's lists.

All stable matchings form a distributive lattice, with M_0 and M_z as its top and bottom. Gusfield and Irving give an elaborate treatment of the structure of this lattice in [7]. We shall see that all matchings achievable by our coalition strategy also form a lattice with M_0 at the bottom. To be more specific, we define a partial order among all matchings that are at least as good as M_0. Matching M, with a set of possible coalitions in it, can be transformed into one of its immediate predecessors in the poset by realizing any one of its coalitions. Moreover, after attaining all the possible coalitions, the resulting matching would be a maximal element in the poset, which satisfies the strong pareto-optimality, i.e., no subset of men can exchange their partners and all improve.

In fact, if the goal is only to find the best lying strategy for men as a whole, then we can directly apply the algorithm for finding the pareto-optimal matching [1]. This algorithm has time complexity of $O(n+m)$ where m is the total length of preference lists truncating the parts following the M_0-partners. But our first two questions are still unanswered. Note that multiple maximal elements may exist, and so finding any one of them does not tell which men are cabalists and which are hopeless.

We point out one more subtlety between the first two questions. After identifying hopeless men, we can safely remove their partners from other men's lists. But the remaining women (who are the partners of cabalists) in their lists are *not* all possible candidates for them. In other words, even if a woman is a partner of some cabalist, it does not follow that all other men putting her on the left of their list can get her by organizing a cabal including themselves.

We now present the algorithm. Create an *envy graph* $G = (\mathcal{M}, E)$, in which a *covet arc* is directed from m to m' if m prefers $M_0(m')$ over $M_0(m)$. Walking along a covet arc from m to m' translates into m being matched to $M_0(m')$. Hence, a directed cycle indicates a possible cabal. A breadth-first-search can check whether a specific man belongs to some cycle. But it is better to find all cabalists in linear time and trim all covet arcs that do not belong to any cycle.

The classical algorithm for finding strongly connected components [14] achieves exactly this goal. We can remove the arcs connecting two strongly connected components since they do not belong to any cycle. For each man, each remaining arc is a possible cabal and a breadth-first-search can be used to find it. Moreover, men who are left without covet arcs are hopeless.

As to finding the maximal element with pareto-optimality, we refer to the top-trading-cycles method [13]. Abraham et al. [1] gave an implementation with linear-time complexity.