Two-Sided Matchings: An Algorithm for Ensuring They Are Minimax and Pareto-Optimal

Steven, Brams and Marc, Kilgour

New York University, Wilfrid Laurier University

7 July 2013

Two-Sided Matchings: An Algorithm for Ensuring They Are Minimax and Pareto-Optimal

Steven J. Brams
Department of Politics
New York University
New York, NY 10012
USA
steven.brams@nyu.edu

D. Marc Kilgour
Department of Mathematics
Wilfrid Laurier University
Waterloo, Ontario N2L 3C5
CANADA
mkilgour@wlu.ca

July 2013

Abstract

Gale and Shapley (1962) proposed the *deferred-acceptance algorithm* for matching (i) college applicants and colleges and (ii) men and women. In the case of the latter, it produces either one or two stable matches whereby no man and woman would prefer to be matched with each other rather than with their present partners. But stable matches can give one or both players in a pair their worst match, whereas the *minimax algorithm* that we propose, which finds all assignments that minimize the maximum rank of players in matches, avoids such assignments. Although minimax matches may not be stable, at least one is always Pareto-optimal: No other matching is at least as good for all the players and better for one or more. If there are multiple minimax matches, we propose criteria for choosing the most desirable among them and also discuss the settings in which minimax matches seem more compelling than deferred-acceptance matches when they differ. Finally, we calculate the probability that minimax matches differ from deferred-acceptance matches in a simple case.

Two-Sided Matchings: An Algorithm for Ensuring They Are Minimax and Pareto-Optimal

1. Introduction

Lloyd S. Shapley and Alvin E. Roth shared the 2012 Nobel Prize in Economic Sciences for developing (Shapley), and extending and applying (Roth), the Gale-Shapley (1962) *deferred-acceptance algorithm* for matching people or other items into stable pairings. In their 1962 paper, Gale and Shapley applied their algorithm to the matching of (i) college applicants and colleges and (ii) men and women. We focus on the latter in this paper, but we mention other applications later.

In section 2, we illustrate the deferred-acceptance algorithm with a simple example, and in section 3 we discuss desirable properties of matchings. In section 4 we present the *minimax algorithm*, a new matching algorithm that minimizes the maximum rank of players in all pairings.

We show that a deferred-acceptance is not always minimax. In fact, the algorithm may produce matchings that are worst for one or both members of a pairing, even though there exist matchings in which no player suffers from such an outcome.

By contrast, the minimax algorithm tends to produce more balanced matchings, but they may not be *stable*—a man and a woman who are matched with different partners may prefer to be matched with each other. But we show in section 5 that at least one minimax matching is always *Pareto-optimal*: No other matching is at least as good for all the players and better for one or more. However, some matchings produced by the minimax algorithm may not be Pareto-optimal.

Stable matchings, in which no pair of players would prefer to switch partners and join each other, are a subset of the Pareto-optimal matchings. The latter also include matchings in which two unmatched players prefer each other, but if they partner, at least one of their former partners must end up worse off.

Whether the stable matchings produced by the deferred-acceptance algorithm are the same or different—in which case they favor either men or women—they tend to be one-sided, whereas minimax matchings, if different from stable matchings, are generally less lopsided. And different they may be: The matchings (either one or two) produced by the deferred-acceptance algorithm may have no overlap with the (one or more) matchings produced by the minimax algorithm, as we show in section 6.

Because there can be multiple minimax matchings, it is useful to have criteria, which we suggest in section 7, to distinguish those that are most desirable. In section 8, we provide some statistics—showing, for example, that about 17 percent of the possible deferred-acceptance matchings of three and three women are not Pareto-optimal minmax matchings—and discuss various extensions and applications of the matchings.

We use examples not only to illustrate the algorithms but also to prove, or illustrate the proofs of, several propositions that generalize the examples. These examples, which provide insight into why the different results arise, highlight circumstances under which Pareto-optimal minimax matchings may be preferred to deferred-acceptance matchings when they diverge.

2. The Deferred-Acceptance Algorithm

Assume that n men have strict preferences over n women (no ties), and that the women likewise have strict preferences over the men. Denote the men by $M = \{m_1, m_2,$

$\ldots, m_n\}$ and the women by $W = \{w_1, w_2, \ldots, w_n\}$; our assumption is that each man m_i, $i = 1, 2, \ldots, n$, has a strict preference order over W, and that each woman w_i, $i = 1, 2, \ldots, n$, has a strict preference order over M. We wish to match each man to a woman and each woman to a man (we assume there are no unacceptable choices). While these and other assumptions of the matching problem have been relaxed over the past 50 years, they have no substantial impact on our main results.[1]

The *deferred-acceptance algorithm* allows for either the men or the women to be the suitors and make offers to members of the opposite gender. We assume that the men are the suitors in stating the algorithm below, but the algorithm could be as well be presented with the women as suitors, making offers to the men.

1. Each man m_i, $i = 1, 2, \ldots, n$, makes an offer to the woman who is his 1st choice.

2. If a woman receives more than one offer, she *holds onto* the one that she most prefers and rejects all the others.

3. Each man whose offer was rejected makes an offer to the next woman in his preference order.

4. Repeat steps 2-3 until there are no further offers to be made because each man has been accepted by a woman to whom he made an offer.

5. Match each man to the woman holding his offer.

The men will cease making offers when, for each man, there is a woman who has not rejected him, so there is no need for him to make an offer to a lower-ranked woman.

[1] Good reviews of the literature on stable matchings and related topics, including both the theory and its applications, are given in Roth (2008), Sönmez and Ünver (2010), and Economic Sciences Prize Committee of the Royal Swedish Academy of Sciences (2012).

At the end, therefore, each man is matched to a different woman, ensuring that the algorithm terminates with n pairs. It is not difficult to show that the maximum number of offers that men can make is $n^2 - 2n + 2$ (Gura and Maschler, 2008, p. 26).

A matching is *stable* iff there is no unmatched man and woman who would each prefer the other to their current partners. Gale and Shapley proved that all deferred-acceptance matchings are stable. This follows from the fact that if man, m_i, prefers to be matched to woman, w_j—rather than the woman he is matched with—he would already have made an offer to w_j. Thus, w_j must have rejected m_i's offer, which she would have done only if she was also holding onto an offer from a man she preferred to m_i. Consequently, w_j must be matched with a man she prefers to m_i. Therefore, after every woman has accepted an offer, no two players who have different partners would prefer to be paired with each other.

Example 1 ($n = 3$)

Suppose three men, m_1, m_2, and m_3, strictly rank three women, w_1, w_2, and w_3, as follows (ignore for now the vertical bars separating the two highest-ranked players from the lowest-ranked player):

m_1: $w_1 > w_2 \mid > w_3$ $\qquad$ w_1: $m_1 > m_2 \mid > m_3$

m_2: $w_1 > w_2 \mid > w_3$ $\qquad$ w_2: $m_1 > m_3 \mid > m_2$

m_3: $w_2 > w_3 \mid > w_1$ $\qquad$ w_3: $m_2 > m_3 \mid > m_1$

We apply the deferred-acceptance algorithm—both the *men-offer* and *women-offer* versions—in this example.

If men-offer, there are three rounds:

1. m_1 and m_2 both make offers to w_1, and m_3 makes an offer to w_2, their highest-ranked women. w_1 rejects the offer that she less prefers (from m_2, whom she ranks 2nd) and holds onto the other offer (from m_1, whom she ranks 1st). w_2 holds onto her only offer (from m_3, whom she ranks 2nd).

2. m_2, who was rejected initially, makes an offer to his next-highest-ranked woman, w_2. Now w_2 holds onto offers from both m_2 and m_3; because she prefers m_3, she rejects the offer from m_2.

3. m_2, who has been rejected again, now makes an offer to w_3, his next-most-preferred woman (but his 3rd choice). Now each woman has exactly one offer (so they must be from different men). There are no more rounds, and the women accept their offers, producing the matching with the rankings shown below:

(1.1) $(m_1, w_1), (m_2, w_3), (m_3, w_2)$: (1st , 1st), (3rd, 1st), (1st, 2nd)

Note that m_2 receives his worst (3rd) choice, whereas no woman receives a choice below 2nd. Nevertheless, there is no case of two unmatched players who prefer each other to their partners, which renders this matching stable.

If women-offer, there are two rounds:

1. w_1 and w_2 both make offers to m_1, and w_3 makes an offer to m_2, their highest-ranked men. m_1 rejects the offer he less prefers (from w_2, whom he ranks 2nd) and holds onto the other offer (from w_1, whom he ranks 1st). m_3 holds onto his only offer (from w_2, whom he ranks 1st).

2. w_2, who was rejected initially, makes an offer to her next-highest-ranked man, m_3. Now each man has exactly one offer. There are no more rounds, and the men accept

the offers they have, resulting in the same matching, (1.1), as obtained in the men-offer algorithm. However, as we will illustrate later, the men-offer and women-offer stable matchings may be different.[2]

When the men-offer algorithm is applied in our example, m_2 is forced to back down first to his 2nd choice (w_2) in step 2, and then to his 3rd choice (w_3) in step 3. This enables m_1 and m_3 to obtain their 1st choices, because both w_1 and w_3 prefer them to m_2.

In the women-offer application, the offers that w_1 and w_3 make in step 1 to m_1 and m_2, respectively, are realized in step 2, whereas w_2's offer to m_1 switches to m_3 in step 2. Unfortunately for m_2, w_2, the only woman to make him an offer, is his 3rd choice, and he must accept her offer when the procedure ends.

Is there another matching in which no person is stuck with a 3rd choice? The answer is "yes," which we show later, but first we define two properties of matchings.

3. Two Properties of Matchings

Consider any one-to-one matching of the set of men, M, and the set of women, W. Define the *depth* of the matching to be the lowest ranking of any player for his or her partner. In Example 1, the depth of the deferred-acceptance is 3, because one player, m_2, is matched with his 3rd choice, w_3.

It is clear that the depth of any matching must be at least 1 and at most n. We define a *minimax matching* to be any matching of minimum depth. In section 4, we show that in Example 1 there is a minimax matching of depth 2.

---

[2] Gura and Maschler (2008, pp. 53-54) give a simple proof that, if the matchings given by the men-offer and women-offer variants of the deferred-acceptance algorithm are identical, then it is the unique stable matching.

Another property of a matching is Pareto-optimality, which we define in terms of Pareto-improvements: One matching is a *Pareto-improvement* on another if the matchings are not identical, and one or more players prefer their partners in the first matching to their partners, if different, in the second, as we illustrate with the following example:

Example 2 (n = 3)

m_1: $w_1 > w_2 > w_3$ w_1: $m_1 > m_2 > m_3$

m_2: $w_2 > w_1 > w_3$ w_2: $m_2 > m_1 > m_3$

m_3: $w_2 > w_3 > w_1$ w_3: $m_3 > m_1 > m_2$

The matching

(2.1) $(m_1, w_1), (m_2, w_2), (m_3, w_3)$: $(1^{st}, 1^{st}), (1^{st}, 1^{st}), (2^{nd}, 1^{st})$

is a Pareto-improvement on (or Pareto-dominates) the matching

(2.2) $(m_1, w_2), (m_2, w_1), (m_3, w_3)$: $(2^{nd}, 2^{nd}), (2^{nd}, 2^{nd}), (2^{nd}, 1^{st})$.

Note that both these matchings have depth 2.

A matching is *Pareto-optimal* iff there are no matchings that are Pareto-improvements on it. In Example 2, there can be no Pareto-improvement on (2.1), because any different matching must give at least one woman a lower choice than 1^{st}, which every woman receives in (2.1). Hence (2.1) is a Pareto-optimal matching, whereas (2.2) is not.

4. The Minimax Algorithm

As before, assume that each member of the set of men, M, strictly ranks each member of the set of women, W, and vice versa. Our goal is to pair each man, m_i, with one member of W and each woman, w_j, with one member of M, so as to create a minimax matching, which minimizes the maximum ranking of any player for his or her partner.

As we will prove shortly (Proposition 1), the application of the minimax algorithm yields all minimax matches. It works by finding an assignment of, say, women to men with the least possible depth, and then verifying that the inverse assignment (of men to women) has at most this same depth. To illustrate our algorithm, we will apply it to Example 1, finding a minimax matching that has depth 2, which is the only minimax matching. By contrast, as we showed earlier, the unique deferred-acceptance matching has depth 3, giving one man (m_2) his 3rd choice (w_3).

To find minimax matches, players successively descend their ranks, going from 1st to lower and lower choices. Put another way, they "fall back" on their preferences, which is why Brams and Kilgour (2001) called a bargaining procedure based on a similar idea—in which all players rank alternatives rather than members of the opposite gender—"fallback bargaining."

In the present application, the men (it could as well be the women) descend in their ranks until there is at least one matching of each man to a different woman. If this matching also yields a matching of the women to different men at the same level or above, the descent stops; otherwise it continues, stopping when *both* men and women can be matched for the first time at the same level.

Minimax Algorithm

Step 1. *Set* $k = 1$.

Step 2. *Determine all one-to-one assignments of men to women so that each man is assigned to a woman he ranks at level k or above. Apply step 3 to every such assignment. If there are no such assignments, increase k by* 1 *and repeat step 2.*

Step 3. *In every one-to-one assignment of men to women in which each man is assigned to a woman he ranks at level k or above, determine whether every woman is assigned to a man she ranks at level k or above. Any such assignment is a minimax matching. If every assignment of men to women at level k fails this condition, increase k by* 1 *and repeat step 2.*

We illustrate the algorithm with Example 1:

Example 1 ($n = 3$)

m_1: $w_1 > w_2 \mid > w_3$ $\quad$ w_1: $m_1 > m_2 \mid > m_3$

m_2: $w_1 > w_2 \mid > w_3$ $\quad$ w_2: $m_1 > m_3 \mid > m_2$

m_3: $w_2 > w_3 \mid > w_1$ $\quad$ w_3: $m_2 > m_3 \mid > m_1$

At level $k = 1$, each man is to be matched with his most preferred woman. Clearly, this is impossible, because two men, m_1 and m_2, rank w_1 highest. Therefore, a level 1 assignment is impossible and we proceed to level 2.

At level $k = 2$, indicated by the vertical bars in the men's rankings, there are two ways to assign different women to (m_1, m_2, m_3): (i) (w_1, w_2, w_3) and (ii) (w_2, w_1, w_3). The next step is to check whether these assignments constitute a level 2 or better matching of the women. In fact, these two assignments give the following matching:

(1.2) $\quad$ $(m_1, w_1), (m_2, w_2), (m_3, w_3)$: ($1^{st}$, 1^{st}), (2^{nd}, 3^{rd}), (2^{nd}, 2^{nd})

(1.3) $(m_1, w_2), (m_2, w_1), (m_3, w_3)$: ($2^{nd}$, 1^{st}), (1^{st}, 2^{nd}), (2^{nd}, 2^{nd}) — Minimax

Notice that for (1.2), w_2 gets only her 3^{rd} choice, so the depth of the matching, 3, is worse than the level for the men. Thus, matching (1.2) is not minimax. However, for matching (1.3), the level of descent of men (level 2) equals the lowest ranking of any woman, so matching (1,3) is minimax. (In fact, it is the unique minimax matching.)

It is instructive to ascertain what the matchings would be if we had started with the women. At level 2, indicated by the vertical bars in the women's ranking of the men, there are two possible assignments of different men to (w_1, w_2, w_3): (i) (m_1, m_3, m_2) and (ii) (m_1, m_2, m_3). Assignment (ii) duplicates assignment (ii) for the men, but assignment (i) for the women gives the following matching:

(1.1) $(m_1, w_1), (m_2, w_3), (m_3, w_2)$: ($1^{st}$, 1^{st}), (3^{rd}, 1^{st}), (1^{st}, 2^{nd}) — Deferred acceptance

Lo and behold, matching (1.1) is the deferred-acceptance matching, which is not minimax because it gives m_2 a 3^{rd} choice. In conclusion, only matching (1.2) is a minimax matching—it minimizes the maximum rankings of both players.

The crucial step in the minimax algorithm is the determination of whether k is large enough that it is possible to assign each man to one of his top k choices. Any such assignment is a *system of distinct representatives*. Hall's Theorem (1935) demonstrates that a men-offer k-level assignment exists if and only if, for every $j = 1, 2, \ldots, n$, every subset of j women intersects the top k choices of at least j men. To illustrate, in Example 1 it is easy to verify that, for $j = 1, 2$, or 3, every subset containing j women overlaps the top $k = 2$ choices of at least j men.

The minimax algorithm determines the minimum value of k such that that Hall's condition (sometimes called the "marriage condition") is satisfied for the men.[3] In addition, it generates all minimax matchings, as we next show, including at least one that is Pareto-optimal (as we will prove in section 5).

Proposition 1. *The minimax algorithm produces all matchings that minimize the maximum rankings of the players.*

Proof. Because there are only finitely many matchings, and each matching has a depth between 1 and n inclusive, there must be at least one matching with minimum depth. If the level of descent, k, is less than this minimax depth, then either it will be impossible to assign a woman to each man, or the matching will fail to have depth at most k because of the preferences of the women. When the matching first reaches the minimax depth, every minimax matching will have occurred, because it must be an assignment of women to men, and men to women, such that no player's partner is ranked below the minimax depth. ■

In section 6, we will give an example (Example 4), with four men and four women, that illustrates how step 3 of the algorithm may require a lower level of descent than that given in Example 1 (i.e., level 2).

Extending Example 1 proves the following:

Proposition 2. *If $n \geq 3$, a deferred-acceptance matching need not be minimax.*

Proof. The rankings for matching (1.3) in Example 1 (deferred-acceptance) prove this proposition for three players. This example can be extended to an example with $n >$

[3] It is noteworthy that a system of distinct representatives can be determined in polynomial time; see Hopcroft and Karp (1973).

3 in which, for $i > 3$, m_i and w_i mutually rank each other 1st, and they rank all additional players of the opposite gender below the three men or three women in Example 1. The original three men and women, who rank each other as before, rank the additional men and women lower. Then the descent of the minimax algorithm will go to level 2, as in Example 1, and give matchings identical to those of Example 1, with all additional players paired according to their subscripts (which occur at level 1).

Similarly, the deferred-acceptance algorithm will give the same result as in Example 1 (i.e., assignment (1.3)) plus the same pairing of the additional men and women. Thus, the deferred-acceptance assignment will differ from the minimax assignment in essentially the same way that it does in Example 1 and, therefore, not be minimax. ■

5. Pareto-Optimality and Minimax

Gale and Shapley (1962) showed that deferred-acceptance assignments are stable; hence, matching (1.1) in Example 1 is stable. Direct verification of this result follows from noting that, with one exception, only one man and one woman does not have his or her 1st-choice partner in matching (1.1). Obviously, a player who already has his or her most preferred partner cannot improve on this matching by finding a different partner.

The only remaining pair is (m_2, w_2), who receive—with their present partners in matching (1.1)—their (3rd, 2nd) choices. If, instead, they were paired with each other, they would receive their (2nd, 3rd) choices. Whereas m_2 would benefit by going from a 3rd to a 2nd choice, w_2 would be hurt by going from a 2nd to a 3rd choice. Consequently, w_2 would not make the switch to (m_2, w_2), rendering matching (1.1) stable.

Now consider minimax matching (1.3), in which there are two men and two women who receive 2nd choices in the three pairings. Pairing up m_1 and w_1 results in 1st choices for each, which they would prefer, rendering matching (1.3) unstable.

Finally, consider matching (1.2), in which there are two men and two women who receive either a 2nd or a 3rd choice in the three pairings. Forming the pair (m_3, w_2) gives the players their (1st, 2nd) choices—a one-step improvement over their present (2nd, 3rd) choices—rendering matching (1.2) unstable.

In summary, of the three matchings we found for Example 1, only the deferred-acceptance matching, (1.1), is stable. The unique minimax matching, (1.3), is not stable, but it is Pareto-optimal, as is the matching (1.2), which is neither stable nor minimax..

To show directly that matching (1.3) is Pareto-optimal, note that it is the only matching of the three given by the minimax algorithm that does not give any player his or her 3rd choice. A matching at least as preferable for all players cannot break up a pair who are each other's 1st and 2nd choices—such as (m_1, w_2) and (m_2, w_1) in Example 1—because any other matching would be worse for the player already getting his or her first choice. This observation implies that any matching at least as preferable as (1.3) must preserve the pairings (m_1, w_2) and (m_2, w_1). Therefore, it must also preserve (m_3, w_3). Thus, there are no Pareto-improvements on (1.3).

Proposition 3. *A matching that Pareto-dominates a minimax matching must also be minimax. In particular, at least one minimax matching must be Pareto-optimal.*

Proof. By Proposition 1 and the nature of the descent process, all minimax matchings must have the same depth. Assume that a minimax matching is Pareto-dominated by another matching. Because every player is at least as well off in the dominating matching, and at least one player is better off, the dominating matching must also be minimax. Moreover, Pareto-dominance is irreflexive and transitive, and there are only a finite number of minimax matchings, so at least one of them must be Pareto-optimal. ■

It is straightforward to show that the two matchings of depth 2 we compared in Example 2 are minimax, one of which, (2.1), Pareto dominates the other, (2.2). This example proves the following:

Proposition 4. *A minimax matching can be Pareto-dominated.*

We note also that matching (2.1) is stable, because it is the unique deferred-acceptance matching. If men-offer, their initial offers are accepted, because each man makes an offer to a different woman. If women-offer, w_1 makes an offer to m_1, and w_2 and w_3 both make offers to m_2; m_2 holds onto his offer from w_2, and w_3 makes a subsequent offer to m_3. Now the three men hold onto offers from different women, which yields again matching (2.1).

Unlike the deferred-acceptance assignment in Example 1, (1.1), the deferred-acceptance matching in (2.1) is minimax, showing that the two algorithms may produce the same output. To show that matching (2.2) is not stable, observe that m_1 and w_1 would both prefer to be paired with each other than to remain with their partners in (2.2).

6. Disagreement between Deferred-Acceptance and Minimax Matchings

Example 1, which we showed extends to larger examples (Proposition 2), demonstrated that the deferred-acceptance algorithm, while always giving a stable outcome, can lead to the worst choice (3^{rd}) for one player (m_2) when the other player in a matching (w_3) receives its best choice (1st). But the deferred-acceptance algorithm can produce something even worse in a pairing:

Proposition 5. *A stable matching can pair a man and a woman so that* <u>*both*</u> *players receive their worst choices, which may not be the case in a minimax matching.*

Proof. Consider the following example, to which we apply the men-offer deferred-acceptance algorithm:

Example 3 ($n = 3$)

m_1: $w_1 > w_2 | > w_3$ w_1: $m_3 > m_2 | > m_1$

m_2: $w_1 > w_2 | > w_3$ w_2: $m_3 > m_1 | > m_2$

m_3: $w_1 > w_3 | > w_2$ w_3: $m_3 > m_1 | > m_2$

1. m_1, m_2, and m_3 all make offers to w_1, their highest-ranked woman. w_1 rejects the offers she less prefers from both m_2 and m_1, holding onto the offer she most prefers from m_3, on which she defers a decision.

2. m_1 and m_2, who were rejected initially, both make offers to w_2, their next-highest-ranked woman. w_2 rejects the offer she less prefers from m_2 and holds onto the offer she more prefers from m_1.

3. m_2, who has been rejected again, now makes an offer to w_3, whom he next most prefers (but ranks 3rd). Now each of the women has exactly one offer from a different. There are no more rounds, and the women accept the offers they have:

(3.1) $(m_1, w_2), (m_2, w_3), (m_3, w_1)$: (2nd, 2nd), (3rd, 3rd), (1st, 1st),

Thus, m_2 and w_3 both receive their worst choices.

In the women-offer algorithm, there are also three rounds, and the matching (3.1) also results. Thus, the deferred-acceptance matching, obtained when either the men-offer or women-offer, gives both members of the pair (m_2, w_3) their worst choices.

The minimax algorithm, when applied to Example 3, yields the unique minimax assignment,

(3.2) $(m_1, w_2), (m_2, w_1), (m_3, w_3)$: $(2^{nd}, 2^{nd}), (1^{st}, 2^{nd}), (2^{nd}, 1^{st})$,

which is different from the unique deferred-acceptance matching, (3.1). Moreover, no player receives his or her worst choice in (3.2).

Example 3 can be embedded in an example with $n > 3$ in which the same phenomenon occurs: The unique stable matching contains a worst-worst pair, whereas no player receives his or her worst choice in a minimax matching. To illustrate, the preferences for $n = 6$ are given by

m_1	$4 > 5 > 6 > 1 > 2 > 3$	w_1	$4 > 5 > 6 > 3 > 2 > 1$
m_2	$4 > 5 > 6 > 1 > 2 > 3$	w_2	$4 > 5 > 6 > 3 > 1 > 2$
m_3	$4 > 5 > 6 > 1 > 3 > 2$	w_3	$4 > 5 > 6 > 3 > 1 > 2$
m_4	$4 > 5 > 6 > 1 > 2 > 3$	w_4	$4 > 5 > 6 > 3 > 2 > 1$
m_5	$5 > 6 > 4 > 1 > 2 > 3$	w_5	$5 > 6 > 4 > 3 > 2 > 1$
m_6	$6 > 4 > 5 > 1 > 2 > 3$	w_6	$6 > 4 > 5 > 3 > 2 > 1$

where for brevity the *m*'s and *w*'s have been dropped from the orderings.

It is easy to verify that the unique stable matching is (3.1), augmented by the pairs $(m_4, w_4), (m_5, w_5), (m_6, w_6)$, and that (3.2), with the same augmentation, is a minimax matching. Again, the deferred-acceptance algorithm, and the unique stable matching, give one pair, (m_2, w_3), their mutually worst choices, whereas there is a minimax matching in which no player receives his or her worst choice. ■

In Example 3, m_2 and w_3 end up with their worst choices (3rd) because each of them prefers two other players to the person he or she is matched with, but these other players do not prefer him or her. Consequently, these other players do not get matched with m_2 or w_3, so m_2 and w_3 get stuck with each other.

Examples 1 and 3 demonstrate our next proposition:

Proposition 6. A *deferred-acceptance matching can be different from any minimax matching*.

But there is not always a disjunction between the deferred-acceptance and minimax matchings, as we saw in the case of (2.1) in Example 2.

Indeed, in the next section we give an example in which there are *two* deferred-acceptance matchings, both of which are not only stable but also minimax. In addition, there are two unstable Pareto-optimal minimax matchings.

7. Criteria for Selecting Preferred Minimax Matchings

Finding all the minimax matchings in our final example shows that step 3 of the minimax algorithm can have bite—the first assignments produced by step 2 do not satisfy the check in step 3, so step 2 must be repeated with a larger value of k.

Example 4 ($n = 4$)

m_1: $w_2 > w_4 \mid > w_1 > w_3$ $\qquad$ w_1: $m_2 > m_1 \mid > m_4 > m_3$

m_2: $w_3 > w_1 \mid > w_4 > w_2$ $\qquad$ w_2: $m_4 > m_3 \mid > m_1 > m_2$

m_3: $w_2 > w_3 \mid > w_1 > w_4$ w_3: $m_1 > m_4 \mid > m_2 > m_3$

m_4: $w_4 > w_1 \mid > w_3 > w_2$ w_4: $m_2 > m_1 \mid > m_4 > m_3$

We begin with the deferred-acceptance algorithm. If men-offer, there are three rounds:

1. m_1 and m_3 both make offers to w_2, while m_2 and m_4 make offers to w_3 and w_4, respectively. Next, w_2 rejects the offer from m_1 (whom she ranks 3rd) and holds onto the offer from m_3, whom she ranks 2nd. w_3 and w_4 hold onto their offers.

2. m_1 makes an offer to his next-highest-ranked woman, w_4, who then rejects her earlier offer from m_4 because she prefers m_1.

3. m_4 makes an offer to his next-highest-ranked woman, w_1, who had no previous offer. Now each woman has exactly one offer, so they accept their offers, producing the matching

(4.1) $(m_1, w_4), (m_2, w_3), (m_3, w_2), (m_4, w_1)$: (2nd, 2nd), (1st, 3rd), (1st, 2nd), (2nd, 3rd)

If women-first, there are four rounds:

1. w_1 and w_4 both make offers to m_2, while w_2 and w_3 make offers to m_4 and m_1, respectively. m_2 rejects the offer from w_4, whom he ranks 3rd in favor of the offer from w_1, whom he ranks 2nd. m_4 and m_1 hold onto their offers.

2. w_4 makes an offer to her next-best man, m_1, who rejects his earlier offer from w_3 because he prefers w_4.

3. w_3 now makes an offer to her next-best man, m_4, who rejects his earlier offer from w_2 because he prefers w_3.

4. w_2 now makes an offer to her next-best man, m_3, who had no previous offer. Now each of the men has exactly one offer, so they accept their offers, producing the matching

(4.2) $(m_1, w_4), (m_2, w_1), (m_3, w_2), (m_4, w_3)$: $(2^{nd}, 2^{nd}), (2^{nd}, 1^{st}), (1^{st}, 2^{nd}), (3^{rd}, 2^{nd})$

Clearly, the men do at least as well, and sometimes better, under men-offer than under women-offer, and vice versa for the women. This is an instance of a general phenomenon: When the men-offer and the women-offer matchings differ, as they do in Example 4, then the former, (4.1), is the best possible stable matching for men, and the latter, (4.2), is the best possible stable matching for women (Roth, 2008; Sönmez and Ünver, 2010; Economic Sciences Prize Committee of the Royal Swedish Academy of Sciences, 2012). This is why men-offer matchings are referred to as *men-optimal*, and women-offer as *women-optimal* (there may be other stable matchings as well).[4]

Now apply the minimax algorithm to Example 4. There are no assignments of M at level $k = 1$, but there are two assignments of (m_1, m_2, m_3, m_4) to W at level $k = 2$: (i) (w_2, w_1, w_3, w_4) and (ii) (w_4, w_3, w_2, w_1). Assignment (ii) produces matching (4.1), the men-optimal deferred-acceptance assignment (M):

(4.3) $(m_1, w_2), (m_2, w_1), (m_3, w_3), (m_4, w_4)$: $(1^{st}, 3^{rd}), (2^{nd}, 1^{st}), (2^{nd}, 4^{th}), (1^{st}, 3^{rd})$

(4.1) $(m_1, w_4), (m_2, w_3), (m_3, w_2), (m_4, w_1)$: $(2^{nd}, 2^{nd}), (1^{st}, 3^{rd}), (1^{st}, 2^{nd}), (2^{nd}, 3^{rd})^*$ — M

Both (4.1) and (4.3) must be rejected as minimax matchings, however, because in each case player is matched to a partner whose rank is below k^{th}.

[4] Gale and Shapley (1962, p. 11) give an example with $n = 3$ in which the unique minimax assignment is different from the two deferred-acceptance assignments but, nevertheless, is stable.

Therefore we set $k = 3$, and consider all possible assignments of $M = \{m_1, m_2, m_3, m_4\}$ to W such that each man is assigned one of his top $k = 3$ women. In addition to the two assignments at level $k = 2$: (i) (w_2, w_1, w_3, w_4) and (ii) (w_4, w_3, w_2, w_1), there are six others: (iii) (w_2, w_3, w_1, w_4), (iv) (w_2, w_4, w_3, w_1), (v) (w_2, w_4, w_1, w_3), (vi) (w_4, w_1, w_2, w_3), (vii) (w_1, w_3, w_2, w_4), and (viii) $(w_1, w_4, w_2\ w_3)$. Thus, we have the following eight candidates for minimax matching, including both the women-optimal (W) and men-optimal (M) matchings:

(4.3) $(m_1, w_2), (m_2, w_1), (m_3, w_3), (m_4, w_4)$: $(1^{st}, 3^{rd}), (2^{nd}, 1^{st}), (2^{nd}, 4^{th}), (1^{st}, 3^{rd})$

(4.1) $(m_1, w_4), (m_2, w_3), (m_3, w_2), (m_4, w_1)$: $(2^{nd}, 2^{nd}), (1^{st}, 3^{rd}), (1^{st}, 2^{nd}), (2^{nd}, 3^{rd})^*$ — M

(4.4) $(m_1, w_2), (m_2, w_3), (m_3, w_1), (m_4, w_4)$: $(1^{st}, 3^{rd}), (1^{st}, 3^{rd}), (3^{rd}, 4^{th}), (1^{st}, 3^{rd})$

(4.5) $(m_1, w_2), (m_2, w_4), (m_3, w_3), (m_4, w_1)$: $(1^{st}, 3^{rd}), (3^{rd}, 1^{st}), (2^{nd}, 4^{th}), (2^{nd}, 3^{rd})$

(4.6) $(m_1, w_2), (m_2, w_4), (m_3, w_1), (m_4, w_3)$: $(1^{st}, 3^{rd}), (3^{rd}, 1^{st}), (3^{rd}, 4^{th}), (3^{rd}, 2^{nd})$

(4.2) $(m_1, w_4), (m_2, w_1), (m_3, w_2), (m_4, w_3)$: $(2^{nd}, 2^{nd}), (2^{nd}, 1^{st}), (1^{st}, 2^{nd}), (3^{rd}, 2^{nd})^*$ — W

(4.7) $(m_1, w_1), (m_2, w_3), (m_3, w_2), (m_4, w_4)$: $(3^{rd}, 2^{nd}), (1^{st}, 3^{rd}), (1^{st}, 2^{nd}), (1^{st}, 3^{rd})^*$

(4.8) $(m_1, w_1), (m_2, w_4), (m_3, w_2), (m_4, w_3)$: $(3^{rd}, 2^{nd}), (3^{rd}, 1^{st}), (1^{st}, 2^{nd}), (3^{rd}, 2^{nd})^*$

We conclude that the minimax matchings for Example 4 are (4.1), (4.2), (4.7), and (4.8). Note that matchings (4.3), (4.4), (4.5), and (4.6) are not minimax, because in each one some player's partner is his or her 4^{th} choice. Moreover, by Proposition 3, all four minimax matchings are Pareto-optimal, because none of them Pareto-dominates any other.

We observed that (4.1) and (4.3) were obtained by searching for assignments of *M* to the sets of top-2 women; similarly, (4.2) and (4.8) could have been found by searching for assignments of *W* to the top-2 men. But note that (4.7) gives at least one 3rd choice to *both* a man and a woman, so this minimax matching could not have been found using a short-cut procedure.

Matchings (4.1) and (4.2), because they are deferred-acceptance matchings, are stable. The other two minimax matchings are not stable: In the case of matching (4.7), if m_1 and w_4 formed a pair, they would get their (2nd, 2nd) choices compared with their present (3rd, 3rd) choices; in the case of matching (4.8), if m_2 and w_1 formed a pair, they would get their (2nd, 1st) choices compared with their present (3rd, 2nd) choices.

The four minimax assignments give varying numbers of 3rd choices to one or both players; they can be ranked from best to worst as follows:

1. One 3rd choice to one gender: (4.2)
2. Two 3rd choices to one gender: (4.1)
3. One 3rd choice to one gender, and two 3rd choices to the other gender: (4.7)
4. Three 3rd choices to one gender: (4.8)

This ranking is based on a (i) *primary criterion* of minimizing the total number of low choices of players, and a (ii) *secondary criterion* that the low choices be split as equally as possible between the genders. We have applied these criteria lexicographically, so the primary criterion gives a ranking based on the number of 3rd choices, and the secondary criterion distinguishes ranks (3) and (4), giving priority to (3) over (4).

The rationale of the primary criterion is that a preferred minimax assignment should minimize the total number of low choices (3rd in Example 4) that players suffer (from one to three in Example 4)—offering relief, insofar as possible, to the worst-off players—which is a criterion championed by Rawls (1971).[5] But if two assignments give the same number of low choices, the one that splits them between the genders is better than one that does not.[6]

Low choices are not only bad for a player who is paired with his of her unappealing partner, but the partner may also be unhappy that he or she is ranked low (should this information be revealed). If one prizes symmetry, mutually low choices of (3^{rd}, 3^{rd}), given by the deferred-acceptance algorithm in Example 3, might actually be better than the (3^{rd}, 1^{st}) and (3^{rd}, 1^{st}) choices in Examples 1 and 4.

On the other hand, the minimax matchings in Examples 1 and 3 give the players, at worst, a 2^{nd} choice, which is presumably less ego-damaging than the 3^{rd} choices given by the deferred-acceptance assignments in these examples. In general, minimax matchings even out, at the highest level attainable, the rankings of players for each other.

This is not always true of matchings that minimize the sum of player ranks, which is what Gusfield and Irving (1989, pp. 128-129) call "egalitarian matches." To be sure, the sums for the minimax matchings in Example 2 and 3, and the most preferred minimax matching in Example 4, (4.2), are the smallest sums for matchings in these examples.

But in Example 1, the deferred-acceptance matching, (1.1), has the smallest sum (9), whereas the minimax matching, (1.3), has sum 10. However, (1.1) gives a pairing of

---

[5] Boudreau and Knoblauch (2013) also analyze Rawlsian-optimal matchings, focusing on their social-welfare implications. They show that players may obtain lower social welfare from stable allocations than from unstable ones, especially when their preferences are correlated, which is what they call the "price of stability." To minimize this price, our minimax algorithm ensures that as few players as possible—individuals as well as pairs—suffer poor matchings.

[6] It is worth pointing out that, in Example 4, the women-optimal deferred-acceptance matching, (4.2), ranks higher than the men-optimal one, (4.1), because it inflicts the players with fewer 3^{rd} choices (one versus two). Thus, if one must choose between these two stable matchings according to our criteria, the women-optimal one is more desirable.

(3rd, 1st) choices, whereas minimax matching (1.3) gives all pairs, at worst, 2nd choices. Because it precludes worst choices, the minimax matching in Example 1 seems to us more egalitarian than the deferred-acceptance matching, even though the minimax matching does not minimize the rank sum.[7]

8. Statistics and Related Results

In the simple case of three men and three women, fix the preferences of, say, m_1, and consider the $(6!)^5 = 7{,}776$ possible preference rankings of the other two men and the three women.[8] All have either one or two deferred-acceptance matchings; 6,488 of these rankings yield at least one deferred-acceptance matching which is also minimax (83.4 percent).

This leaves 7,776 – 6,488 = 1,288 (16.6 percent) of the preference rankings in which the minimax matchings differ from the deferred-acceptance matchings. Of these, 1,008 (78.2 percent) have no stable minimax matching, though by Proposition 2 there must be at least one Pareto-optimal minimax matching. These constitute 12.9 percent of all rankings. In our view, these Pareto-optimal minimax matchings, even though they are not stable, deserve consideration as desirable matchings, as do stable minimax matchings that are not deferred-acceptance matchings (3.6 percent).

We conjecture that the proportion of minimax assignments that are not deferred-acceptance assignments increases rapidly with the number of players, approaching 1 as

---

[7] Knuth (1997, pp. 50-51) reports that Stan Selkow proposed an algorithm which starts from the men-optimal and the women-optimal stable matches and progressively shifts them toward those that reduce the distance of the most unhappy player from his or her most preferred stable match. Gusfield and Irving (pp. 135-143) propose a similar approach with "parametric stable marriages." But these algorithms are not applicable when the men-optimal and women-optimal matches are the same, but not minimax, as in Examples 1 and 3.

[8] We thank Eli H. Ross for making the computer calculations on which these statistics are based.

the number of players approaches infinity. If true, this would strengthen our case that minimax assignments are a compelling alternative to deferred-acceptance matchings.

There have been many extensions and generalizations of matching algorithms, often motivated by applications. They include the following:

1. *Different numbers of men and women*. If there are, say, more men than women, then the algorithms run until all the women are matched, leaving the remaining men unmatched. (Women are advantaged in this case by there being fewer of them.)

2. *Incomplete or nonstrict preferences, and unacceptable pairings*. If one side of a match does not rank all members of the other side, or if one side considers certain members of the other side unacceptable, these pairings will not be made. If preferences are not strict, more minimax matches become possible.

3. *College admissions and hospital residencies*. Because different colleges (hospitals) admit different numbers of students (residents), the algorithm must be adjusted. Each college (hospital) is duplicated once for each of its openings, with each copy having the same preference for the students (residents). Similarly, each student (resident) has the same preference for each opening of a college (hospital).

4. *Organ matches*. Because preference is defined by medical compatibility, he problem of matching organ donors and recipients is symmetric.

5. *Roommates*. Roommates rank each other rather than members of another group. In this problem, a stable matching may not exist (for a simple example, see Gale

and Shapley, 1962, p. 12), but the minimax algorithm can still be applied and a Pareto-optimal match found.[9]

Pareto-optimal minimax matchings offer more balanced pairings than stable deferred-acceptance matchings that are not minimax, whether the latter are unique or men-optimal and women-optimal. If the matches are administered by a central clearinghouse, as are the National Medical Residency Program and school-choice matching programs in Boston and New York City (Roth 2008; Sönmex and Ünver, 2011), minimax matchings that are Pareto-optimal will be difficult to manipulate.[10]

But in a smaller setting—in which, for example, a man and a woman might discover that they were matched with inferior choices—it may be better to preserve stability by using the deferred-acceptance algorithm. In larger and more anonymous settings, however, the minimax algorithm, by closing the distance in matched choices, is appealing.

Even when the deferred-acceptance algorithm produces only one stable matching, it may give one or both players in a pair a worst choice when a minimax matching does not (as in Examples 1 and 3). If there are multiple minimax matchings, the primary and secondary criteria we suggested provide a guide for selecting more balanced matches.

That politics is the art of compromise is a cliché, but in the matching of people (men and women), people and institutions (colleges, hospitals), or people and the organs that they need (e.g., kidneys), compromises that avoid lopsided matches are desirable. In

---

[9] When, as in the roommates problem, each player ranks every other player, Brams, Jones, and Kilgour (2005) define “stable” and “semi-stable” coalitions, which are not restricted to pairs but may contain any number up to, and including, all players.

[10] Much research, using game theory, complexity theory, and optimization methods, has been applied to strategic questions, which are discussed in several of the aforementioned references.

our opinion, the minimax algorithm facilitates the search for such compromises, turning an art into more of a science.

References

Boudreau, James W., and Vicki Knoblauch (2013). "Preferences and the Price of Stability in Matching Markets." *Theory and Decision* 74, no. 4 (December): 565-589.

Brams, Steven J., Michael A. Jones, and D. Marc Kilgour, and (2005). "Forming Stable Coalitions: The Process Matters." *Public Choice* 125, nos. 1 and 2 (October): 67-94.

Brams, Steven J., and D. Marc Kilgour (2001). "Fallback Bargaining." *Group Decision and Negotiation* 10, no. 4 (July): 287-316.

Economic Sciences Prize Committee of the Royal Swedish Academy of Sciences (2012). "Stable Allocations and the Practice of Market Design." Stockholm, Sweden. http://www.nobelprize.org/nobel_prizes/economics/laureates/2012/advanced-economicsciences2012.pdf

Gale, D., and Lloyd S. Shapley (1962). "College Admissions and the Stability of Marriage." *American Mathematical Monthly* 69, no. 1 (January): 9-15.

Gura, Ein-Ya, and Michael Maschler (2008). *Insights into Game Theory: An Alternative Mathematical Experience*. Cambridge, UK: Cambridge University Press.

Gusfield, Dan, and Robert W. Irving (1989). *The Stable Marriage Problem: Structure and Algorithms*. Cambridge, MA: MIT Press.

Hall, Philip (1935). "On Representatives of Sets." *Journal of the London Mathematical Society* 10, no. 1: 26-30.

Hopcroft, John E., and Richard M. Karp (1973). "An $n^{5/2}$ Algorithm for Maximum Matchings in Bipartite Graphs," *SIAM journal of Computing* 2, no. 4: 225-231.

Knuth, Donald E. (1997). *Stable Marriage and Its Relation to Other Combinatorial Problems: An Introduction to the Mathematical Analysis of Algorithms*. Providence, RI: American Mathematical Society.

Rawls, John (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Roth, Alvin E. (2008). "Deferred Acceptance Algorithms: History, Theory, Practice, and Open Questions." Internat*ional Journal of Game Theory* 36: 537-569.

Roth, Alvin E., and Marilda A. Oliveira Sotomayor (1990). *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge, UK: Cambridge University Press.

Sönmez, Tayfun, and M. Utku Ünver (2011). "Matching, Allocation, and Exchange of Discrete Resources," in Jess Benhabib, Albert Bison, and Matthew O. Jackson (eds.), *Handbook of Social Economics*, Vol. 1A. Amsterdam: North Holland, pp. 781-852.