Two-Sided Markets: Stable Matching
(DRAFT: Not for distribution)

Federico Echenique[*]

Nicole Immorlica[†]

Vijay V. Vazirani[‡]

1 Introduction

The field of matching markets was initiated by the seminal work of Gale and Shapley on stable matching. Stable matchings have remarkably deep and pristine structural properties, which have led to polynomial time algorithms for numerous computational problems as well as quintessential game-theoretic properties. In turn, these have opened up the use of stable matchings to a host of important applications.

This chapter[1] will deal with the following four aspects:

1. Gale and Shapley's Deferred Acceptance Algorithm for computing a stable matching; we will sometimes refer to it as the DA Algorithm.
2. Incentive compatibility properties of this algorithm.
3. The fact that the set of all stable matchings of an instance forms a finite, distributive lattice, and the rich collection of structural properties associated with this fact.
4. Linear programing approach to computing stable matchings.

A general setting: A setting of the stable matching problem which is particularly useful in applications is the following (this definition is quite complicated because of its generality, and can be skipped on the first reading).

Definition 1. Let W be a set of n workers and F a set of m firms. Let c be a *capacity function* $c : F \to \mathbb{Z}_+$ giving the maximum number of workers that can be matched to a firm; each worker can be matched to at most one firm. Also, let $G = (W, F, E)$ be a bipartite graph on vertex sets W, F and edge set E. For a vertex v in G, let $N(v)$ denote the set of its neighbors in G. Each

[*]Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125.

[†]Microsoft Research Lab – New York City, New York, NY 10011.

[‡]Department of Computer Science, University of California at Irvine, Irvine, CA 92697.

[1]This chapter will appear in:
Online and Matching-Based Market Design. Federico Echenique, Nicole Immorlica and Vijay V. Vazirani, Editors. Cambridge University Press. © 2021

worker w provides a strict preference list $l(w)$ over the set $N(w)$ and each firm f provides a strict preference list $l(f)$ over the set $N(f)$. We will adopt the convention that each worker and firm prefers getting matched to one of its neighbors to remaining unmatched, and it prefers remaining unmatched to getting matched to a non-neighbor[2]. If a worker or firm remains unmatched, we will say that it is matched to $\perp$

We wish to study all four aspects stated for this setting. However, it will be quite unwise and needlessly cumbersome to study the aspects directly in this setting. It turns out that the stable matching problem offers a natural progression of settings, hence allowing us to study the aspects gradually in increasing generality.

1. **Setting I:** Under this setting $n = m$, the capacity of each firm is one and graph G is a complete bipartite graph. Thus in this setting each worker and firm has a total order over the other side. This simple setting will be used for introducing the core ideas.
2. **Setting II:** Under this setting n and m are not required to be equal and G is arbitrary; however, the capacity of each firm is still one. The definition of stability becomes more elaborate, hence making all four aspects more difficult in this setting. Relying on the foundation laid In Setting I, we will present only the additional ideas needed.
3. **Setting III:** This is the general setting defined above. We will give a reduction from this setting to Setting II, so that the algorithm and its consequences carry over without additional work.

2 The Gale-Shapley Deferred Acceptance Algorithm

In this section we will define the notion of a stable matching for all three settings and give an efficient algorithm for finding it.

2.1 The DA Algorithm for setting I

In this setting, the number of workers and firms is equal, i.e., $n = m$, and each firm has unit capacity. Furthermore, each worker and each firm has a total order over the other side.

Notation: If *worker w prefers firm f to f'*, then we represent this as $f \succ_w f'$; a similar notation is used for describing the preferences of a firm.

We next recall a key definition from graph theory. Let $G = (W, F, E)$ be a graph with equal numbers of workers and firms, i.e., $|W| = |F|$. Then, $\mu \subseteq E$ is a *perfect matching* in G if each vertex of G has exactly one edge of μ incident at it. If so, μ can also be viewed as a bijection between W to F. If $(w, f) \in \mu$ then we will say that *μ matches w to f* and use the notation $\mu(w) = f$ and $\mu(f) = w$.

Definition 2. Worker w and firm f form a *blocking pair* with respect to a perfect matching μ, if they prefer each other over their partners in μ, i.e., $w \prec_f \mu(f)$ and $f \prec_w \mu(w)$.

[2]An alternative way of defining preference lists, which we will use in Section 3.2 is the following: Each worker w has a preference list over $F \cup \{\perp\}$, with firms in $N(w)$ listed in the preference order of w, followed by $\perp$, followed by $(F \setminus N(w))$ listed in arbitrary order. Similarly, each firm f's preference list is over $W \cup \{\perp\}$.

If (w, f) form a blocking pair with respect to perfect matching μ, then they have incentive to secede from matching μ and pair up by themselves. The significance of the notion of stable matching, defined next, is that no worker-firm pair has an incentive to secede from this matching. Hence such matchings lie in the *core* of the particular instance; this key notion will be introduced in Chapter ??. For now, recall from cooperative game theory that the core consists of solutions under which no subset of the agents can gain more (i.e., with each one gaining at least as much and at least one agent gaining strictly more) by seceding from the grand coalition. Additionally, in Chapter ?? we will also establish that stable matchings are efficient and individually rational.

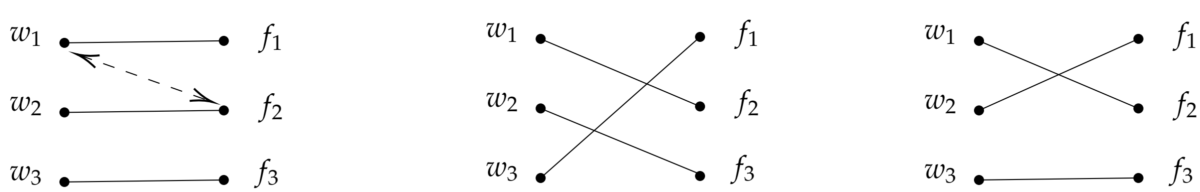
Definition 3. A perfect matching μ with no blocking pairs is called a *stable matching*.

It turns out that every instance of the stable matching problem with complete preference lists has at least one stable matching. Interestingly enough, this fact follows as a corollary of the Deferred Acceptance Algorithm, which finds in polynomial time one stable matching among the $n!$ possible perfect matchings in G.

Example 4. Let I be an instance of the stable matching problem with 3 workers and 3 firms and the following preference lists:

$w_1 : f_2, f_1, f_3$ $\quad$ $f_1 : w_1, w_2, w_3$
$w_2 : f_2, f_3, f_1$ $\quad$ $f_2 : w_1, w_2, w_3$
$w_3 : f_1, f_2, f_3$ $\quad$ $f_3 : w_1, w_3, w_2$

The next figure shows three perfect matchings in instance I. The first matching is unstable, with blocking pair (w_1, f_2), and the last two are stable (this statement is worth verifying).



(a) An unstable perfect matching $\quad$ (b) Stable matching 1 $\quad$ (c) Stable matching 2

We next present the Deferred Acceptance Algorithm[3] for Setting I, described in Algorithm 8. The algorithm operates iteratively, with one side proposing and the other side acting on the proposals received. We will assume that workers propose to firms. The initialization involves each worker marking each firm in its preference list as *uncrossed*.

Each iteration consists of three steps: First, each worker proposes to the best uncrossed firm on its list. Second, each firm that got proposals tentatively accepts the best proposal it received and rejects all other proposals. Third, each worker who was rejected by a firm crosses that firm off its list. If in an iteration each firm receives a proposal, we have a perfect matching, say μ, and the algorithm terminates.

The following observations will lead to a proof of correctness and running time.

[3] The reason for this name is provided in Remark 11.

Algorithm 8. (Deferred Acceptance Algorithm)

Until all firms receive a proposal, do:

1. $\forall w \in W$: w proposes to its best uncrossed firm.
2. $\forall f \in F$: f tentatively accepts its best proposal and rejects the rest.
3. $\forall w \in W$: If w got rejected by firm f, it crosses f off its list.

Output the perfect matching, and call it μ.

Observation 5. *If a firm gets a proposal in a certain iteration, it will keep getting at least one proposal in all subsequent iterations.*

Observation 6. *As the iterations proceed, for each firm, the following holds: once it receives a proposal, it tentatively accepts a proposal from the same or a better worker, according to its preference list.*

Lemma 7. *Algorithm 8 terminates in at most n^2 iterations.*

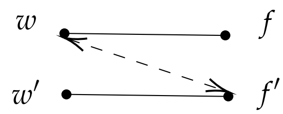
Proof. In every iteration, other than the last one, at least one worker will cross a firm off its preference list. Consider iteration number $n^2 - n + 1$, assuming the algorithm has not terminated so far. Since the total size of the n preference lists is n^2, there is a worker, say w, who will propose to the last firm on its list in this iteration. Therefore by this iteration w has proposed to every firm and every firm has received a proposal. Hence, by Observation 5, in this iteration every firm will get a proposal and the algorithm will terminate with a perfect matching. □

Example 9. The figures below show the two iterations executed by Algorithm 8 on the instance of Example 4. In the first iteration, w_2 will get rejected by f_2 and will cross it from its list. In the second iteration, w_2 will propose to f_3, resulting in a perfect matching.

(a) Iteration 1 (b) Iteration 2

Theorem 10. *The perfect matching found by the DA Algorithm is stable.*

Proof. For the sake of contradiction assume that μ is not stable and let (w, f') be a blocking pair. Assume that $\mu(w) = f$ and $\mu(f') = w'$ as shown in the figure below. Since (w, f') is a blocking pair, w prefers f' to f and therefore must have proposed to f' and got rejected in some iteration, say i, before eventually proposing to f. In iteration i, f' must have tentatively accepted the proposal from a worker it likes better than w. Therefore by Observation 6, at the termination of the algorithm, $w' \succ_{f'} w$. This contradicts the assumption that (w, f') is a blocking pair. □



Blocking pair (w, f')

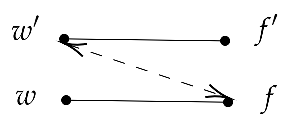
Remark 11. The Gale-Shapley algorithm is called the *Deferred Acceptance Algorithm* because firms do not immediately accept proposals received by them – they defer them and accept only at the end of the algorithm when a perfect matching is found. In contrast, under the Immediate Acceptance Algorithm, each firm immediately accepts the best of the proposal it received; see Chapter ??.

Our next goal is to prove that the DA Algorithm, with workers proposing, leads to a matching that is favorable for workers and unfavorable for firms. We first formalize the terms "favorable" and "unfavorable".

Definition 12. Let S be the set of all stable matchings over (W, F). For each worker w, the *realm of possibilities* $R(w)$ is the set of all firms that w is matched to in S, i.e., $R(w) = \{f \mid \exists \mu \in S \text{ s.t. } (w, f) \in \mu\}$. The *optimal firm* for w is the best firm in $R(w)$ with respect to w's preference list; it will be denoted by optimal(w). The *pessimal firm* for w is the worst firm in $R(w)$ with respect to w's preference list and will be denoted by pessimal(w). The definitions of these terms for firms are analogous.

Lemma 13. *Two workers cannot have the same optimal firm, i.e., each worker has a unique optimal firm.*

Proof. Suppose not and suppose two workers w and w' have the same optimal firm, f. Assume w.l.o.g. that f prefers w' to w. Let μ be a stable matching such that $(w, f) \in \mu$ and let f' be the firm matched to w' in μ. Since $f = \text{optimal}(w')$ and w' is matched to f' in a stable matching, it must be the case that $f \succ_{w'} f'$. Then (w', f) forms a blocking pair with respect to μ, leading to a contradiction. □



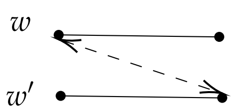
Blocking pair (w', f) with respect to μ.

Corollary 14. *Matching each worker to its optimal firm results in a perfect matching, say μ_W.*

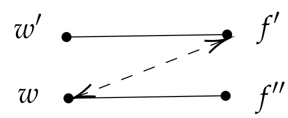
Lemma 15. *Matching μ_W is stable.*

Proof. Suppose not and let (w, f') be a blocking pair with respect to μ_W, where $(w, f), (w', f') \in \mu_W$. Then $f' \succ_w f$ and $w \succ_{f'} w'$.

Since $opt(w') = f'$, there is a stable matching, say μ', s.t. $(w', f') \in \mu'$. Assume that w is matched to firm f'' in μ'. Now since $opt(w) = f$, $f \succ_w f''$. This together with $f' \succ_w f$ gives $f' \succ_w f''$. Then (w, f') is a blocking pair with respect to μ', giving a contradiction. □



(a) Blocking pair (w, f') with respect to μ_W



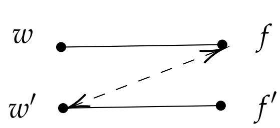
(b) Blocking pair (w, f') with respect to μ'

Proofs similar to that of Lemma 13 and Lemma 15 show that each worker has a unique pessimal firm and the perfect matching that matches each worker to its pessimal firm is also stable.

Definition 16. The perfect matching that matches each worker to its optimal (pessimal) firm is called the *worker optimal (pessimal) stable matching*. The notions of *firm optimal (pessimal) stable matching* are analogous. The worker and firm optimal stable matchings will be denoted by μ_W and μ_F, respectively.

Theorem 17. *The worker-proposing DA Algorithm finds the worker-optimal stable matching.*

Proof. Suppose not, then there must be a worker who is rejected by its optimal firm before proposing to a firm it prefers less. Consider the first iteration in which a worker, say w, is rejected by its optimal firm, say f. Let w' be the worker firm f tentatively accepts in this iteration; clearly, $w' \succ_f w$. By Lemma 13, optimal$(w') \neq f$ and by the assumption made in the first sentence, w' has not yet been rejected by its optimal firm (and perhaps never will be). Therefore, w' has not yet proposed to its optimal firm; let the latter be f'. Since w' has already proposed to f, we have that $f \succ_{w'} f'$. Now consider the worker optimal stable matching μ; clearly, $(w, f), (w', f') \in \mu$. Then (w', f) is a blocking pair with respect to μ, giving a contradiction. □



Blocking pair (w', f) with respect to μ

Lemma 18. *The worker-optimal stable matching is also firm pessimal.*

Proof. Let μ be the worker optimal stable matching and suppose that it is not firm pessimal. Let μ' be the latter stable matching. Now for some $(w, f) \in \mu$, pessimal$(f) \neq w$. Let pessimal$(f) = w'$; clearly, $w \succ_f w'$. Let $w =$ pessimal(f'), then $(w, f'), (w', f) \in \mu'$. Since optimal$(w) = f$ and w is matched to f' in a stable matching, $f \succ_w f'$. Then (w, f) forms a blocking pair with respect to μ', giving a contradiction. □

2.2 Extension to Setting II

Recall that in this setting, each worker and firm has a total preference order over only its neighbors in the graph $G = (W, F, E)$ and $\perp$, with $\perp$ being the least preferred element in each list; matching a worker or firm to $\perp$ is equivalent to leaving it unmatched.

In this setting, a stable matching may not be a perfect matching in G even if the number of workers and firms is equal; however, it will be a maximal matching. Recall that matching $\mu \subseteq E$ is *maximal* if it cannot be extended with an edge from $E - \mu$. As a result of these changes, the definition of stability also needs to be enhanced.

Definition 19. Let μ be any maximal matching in $G = (W, F, E)$. Then the pair (w, f) forms a *blocking pair* with respect to μ if $(w, f) \in E$ and either:

- **Type 1:** w, f are both matched in μ and prefer each other to their partners in μ.
- **Type 2a:** w is matched to f', f is unmatched and $f \succ_w f'$.
- **Type 2b:** w is unmatched, f is matched to w' and $w \succ_f w'$.

Observe that since $(w, f) \in E$, w and f prefer each other to remaining unmatched. Therefore they both cannot be unmatched in μ — this follows from the maximality of the matching.

The only modification needed to Algorithm 8 is to the termination condition, which is: Every worker is either tentatively accepted by a firm or has crossed off all firms from its list. When this condition is reached, each worker of the first type is matched to the firm that tentatively accepted it and the rest remain unmatched. Let μ denote this matching. We will still call this the Deferred Acceptance Algorithm. It is easy to see that Observations 5 and 6 still hold and that Lemma 7 holds with a bound of nm on the number of iterations.

Lemma 20. *The Deferred Acceptance Algorithm outputs a maximal matching in G.*

Proof. Assume that $(w, f) \in E$ and yet worker w and firm f are both unmatched in the matching found by the algorithm. During the algorithm, w must have proposed to f and got rejected. Now, by Observation 5, f must be matched, giving a contradiction. □

Theorem 21. *The maximal matching found by the Deferred Acceptance Algorithm is stable.*

Proof. We need to prove that both types of blocking pairs do not exist with respect to μ. For the first type, the proof is identical to that in Theorem 10 and is omitted.

Assume that (w, f) is a blocking pair of the second type. There are two cases:

Case 1: w is matched, f is not, and w prefers f to its match, say f'. Clearly w will propose to f before proposing to f'. Now, by Observation 5, f must be matched in μ, giving a contradiction.

Case 2: f is matched, w is not, and f prefers w to its match, say w'. Clearly w will propose to f during the algorithm. Since f prefers w to w', it will not reject w in favor of w', hence giving a contradiction. □

Notation: If worker w or firm f is unmatched in μ, then we will denote it as $\mu(w) = \perp$ or $\mu(f) = \perp$. We will denote the set of workers and firms matched under μ by $W(\mu)$ and $F(\mu)$, respectively.

Several of the definitions and facts given in Setting I carry over with small modifications; we summarize these next. The definition of *realm of possibilities* of workers and firms remains the same as before; however, note that in Setting II, some of these sets could be the singleton set $\{\perp\}$. The definition of *optimal and pessimal firm for a worker* also remains the same, with the change that it will be $\perp$ if the realm of possibilities is the set $\{\perp\}$. Let $W' \subseteq W$ be the set of workers whose realm of possibilities is non-empty. Then, via a proof similar to that of Lemma 13, it is easy to see that two workers in W' cannot have the same optimal firm, i.e., every worker in W' has a unique optimal firm.

Next, match each worker in W' to its optimal firm, leaving the remaining workers unmatched. This is defined to be the *worker optimal matching*; we will denote it by μ_W. Similarly, define the *firm optimal matching*; this will be denoted by μ_F. Using ideas from the proof of Lemma 15, it is easy to show that the worker optimal matching is stable. Furthermore, using Theorem 17 one can show that the Deferred Acceptance Algorithm finds this matching. Finally, using Lemma 18, one can show that the worker optimal stable matching is also firm pessimal.

Lemma 22. *The number of workers and firms matched in all stable matchings is the same.*

Proof. Each worker w prefers getting matched to one of the firms that is its neighbor in G over remaining unmatched. Therefore, all workers who are unmatched in μ_W will be unmatched in all other stable matchings as well. Hence for an arbitrary stable matching μ we have $W(\mu_W) \supseteq W(\mu) \supseteq W(\mu_F)$.
Thus $|W(\mu_W)| \geq |W(\mu)| \geq |W(\mu_F)|$. A similar statement for firms is $|F(\mu_W)| \leq |F(\mu)| \leq |F(\mu_F)|$. Since the number of workers and firms matched in any stable matching is equal, $|W(\mu_W)| = |F(\mu_W)|$ and $|W(\mu_F)| = |F(\mu_F)|$. Therefore the cardinalities of all sets given above are equal, hence establishing the lemma. □

Finally, we present the *Rural Hospital Theorem*[4] for Setting II.

Theorem 23. *The set of workers matched is the same under all stable matchings; similarly for firms.*

[4] The name of this theorem has its origins in the application of stable matching to the problem of matching residents to hospitals. The full scope of the explanation given next is best seen in the context of the extension of this theorem to Setting III, given in Section 2.3. In this application it was found that certain hospitals got very poor matches and even remained under-filled; moreover, this persisted even when a hospital-optimal stable matching was resorted to. It turned out that these unsatisfied hospitals were mostly in rural areas and were preferred least by most residents. The question arose if there was a "better" way of finding an allocation. The Rural Hospital Theorem clarified that *every* stable matching would treat under-filled hospitals in the same way, i.e., give the same allocation.

Proof. As observed in the proof of Lemma 22, $W(\mu_W) \supseteq W(\mu) \supseteq W(\mu_F)$. By Lemma 22 these sets are of equal cardinality. Hence they must all be the same set as well. □

2.3 Reduction from Setting III to Setting II

We will first give a definition of blocking pair that is appropriate for Setting III. We will then give a reduction from this setting to Setting II, thereby allowing us to carry over the algorithm and its consequences to this setting directly. Finally, we will prove the Rural Hospital Theorem for Setting III.

Definition 24. Given a graph $G = (V, E)$ and an upper bound function $b : V \to \mathbb{Z}_+$, a set $\mu \subseteq E$ is a *b-matching* if the number of edges of μ incident at each vertex $v \in V$ is at most $b(v)$. Furthermore, μ is a *maximal b-matching* if μ cannot be extended to a valid b-matching by adding an edge from $E - \mu$.

In Setting III, firms have capacities given by $c : F \to \mathbb{Z}_+$. For the graph $G = (W, F, E)$ specified in the given instance in Setting III, define upper bound function $b : W \cup F \to \mathbb{Z}_+$ as follows: for $w \in W$, $b(w) = 1$ and for $f \in F$, $b(f) = c(f)$. Let μ be a maximal b-matching in $G = (W, F, E)$ with upper bound function b. We will say that firm f is *matched to capacity* if the number of workers matched to f is exactly $c(f)$ and it is *not matched to capacity* if f is matched to fewer than $c(f)$ workers. Furthermore, if a set $S \subseteq W$ of workers is matched to firm f under μ, with $|S| \leq c(f)$, then we will use the notation $\mu(f) = S$ and for each $w \in S$, $\mu(w) = f$.

Definition 25. Let μ be a maximal b-matching in $G = (W, F, E)$ with upper bound function b. For $w \in W$ and $f \in F$, (w, f) forms a *blocking pair* with respect to μ if $(w, f) \in E$ and one of the following hold:

- **Type 1:** f is matched to capacity, w is matched to f' and there is a worker w' that is matched to f such that $w \succ_f w'$ and $f \succ_w f'$.
- **Type 2a:** f is not matched to capacity, w is matched to f' and $f \succ_w f'$.
- **Type 2b:** w is unmatched, w' is matched to f and $w \succ_f w'$.

Reduction to Setting II: Given an instance I of Setting III, we show below how to reduce it in polynomial time to an instance I' of Setting II so that there is bijection ϕ between the sets of stable matchings of I and I' such that ϕ and ϕ^{-1} can be computed in polynomial time.

Let I be given by (W, F, E, c) together with preference lists $l(w), \forall w \in W$ and $l(f), \forall f \in F$. Instance I' will be given by (W', F', E') together with preference lists $l'(w), \forall w \in W'$ and $l'(f), \forall f \in F'$, where:

- $W' = W$.
- $F' = \cup_{f \in F} \{f^{(1)}, \ldots, f^{(c(f))}\}$, i.e., corresponding to firm $f \in I$, I' will have $c(f)$ firms, namely $f^{(1)}, \ldots, f^{(c(f))}$.
- Corresponding to each edge $(w, f) \in E$, E' has edges $(w, f^{(i)})$, for each $i \in [1 \ldots c(f)]$.
- $\forall w \in W'$, $l'(w)$ is obtained by replacing each firm, say f, in $l(w)$ by the ordered list $f^{(1)}, ..., f^{(c(f))}$. More formally, if $f \succ_w f'$ then for all $1 \leq i \leq c(i)$ and $1 \leq j \leq c(j)$ we

have $f^{(i)} \succ_w f'^{(j)}$ and for all $1 \leq i < j \leq c(i)$ we have $f^{(i)} \succ_w f^{(j)}$.

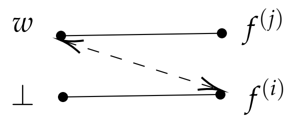
- $\forall f \in F$ and $i \in [1 \ldots c(f)]$, $l'(f^{(i)})$ is the same as $l(f)$.

Lemma 26. *Let μ be a stable matching for instance I of Setting III. Then the following hold:*

- *If firm f is matched to $k < c(f)$ workers, then $f^{(1)}...f^{(k)}$ must be matched and $f^{(k+1)}...f^{(c(f))}$ must remain unmatched.*
- *If $(f^{(i)}, w), (f^{(j)}, w') \in \mu$ with $i < j$, then $w \succ_f w'$.*

Proof. For contradiction assume that $f^{(i)}$ is unmatched and $f^{(j)}$ is matched, to w say, in μ, where $i < j$. Clearly, $f^{(i)} \succ_w f^{(j)}$ and $w \succ_{f^{(i)}} \perp$. Therefore $(w, f^{(i)})$ is a blocking pair; see figure below.

The second proof is analogous, with $\perp$ replaced by w'. □



Blocking pair $\left(w,\ f^{(i)}\right)$

Theorem 27. *There is a bijection between the sets of stable matchings of I and I'.*

Proof. We will first define a mapping ϕ from the first set to the second and then we will prove that it is a bijection.

Let μ be a stable matching of instance I. Assume that set $S \subseteq W$ of workers is matched in μ to firm f and worker $w \in S$ and w is the i^{th} most preferred worker in S with respect to $l(f)$. Then, under $\phi(\mu)$ we will match w to $f^{(i)}$. This defines $\phi(\mu)$ completely.

For contradiction assume that $\phi(\mu)$ is not stable and let $(w, f^{(i)})$ be a blocking pair with respect to $\phi(\mu)$. Assume that $f^{(i)}$ is matched to w', where either $w' \in W'$ or $w' = \perp$. Clearly $f^{(i)}$ prefers w to w', therefore by construction, if w is matched to $f^{(j)}$, then $j < i$, contradicting the fact that $(w, f^{(i)})$ is a blocking pair. Hence w is either unmatched or is matched to $f'^{(k)}$ for some $f' \neq f$ and some k.

In the first case, under μ, w is unmatched, w' is matched to f, $w \succ_f w'$ and $f \succ_w \perp$. In the second case, under μ, w is matched to f', w' is matched to f, $w \succ_f w'$ and $f \succ_w f'$. Therefore in both cases (w, f) is a blocking pair with respect to μ, giving a contradiction.

Finally we observe that ϕ has an inverse map such that $\phi^{-1}(\phi(\mu)) = \mu$. Let μ' be a stable matching of instance I'. If $\mu'(w) = f^{(i)}$, then $\phi^{-1}(\mu')$ matches w to f. The stability of $\phi^{-1}(\mu')$ is easily shown; in particular, because each f has the same preference lists as $f^{(i)}$ for $i \in [1, \ldots, c(f)]$. □

As a consequence of Theorem 27, we can transform the given instance I to an instance I' of Setting II, run the Deferred Acceptance Algorithm on it and transform the solution back to obtain a stable matching for I. Clearly, all notions established in Setting II following the algorithm, such as realm of possibilities and worker optimal and firm optimal stable matching, also carry over to the current setting.

Finally, we present the Rural Hospital Theorem for Setting III.

Theorem 28. *The following hold for an instance in Setting III:*

1. *Over all the stable matchings of the given instance: the set of matched workers is the same and the number workers matched to each firm is also the same.*
2. *Assume that firm f is not matched to capacity in the stable matchings. Then, the set of workers matched to f is the same over all stable matchings.*

Proof. 1). Let us reduce the given instance, say I, to an instance I' in Setting II and apply Theorem 23. Then we get that the set of matched workers is the same over all the stable matchings of I', hence yielding the same statement for I as well. We also get that the set of matched firms is the same over all the stable matchings of I'. Applying bijection ϕ^{-1} to I' we get that the number workers matched to each firm is also the same over all stable matchings.

2). Let μ_W and μ_F be the worker-optimal and firm-optimal stable matchings for instance I, respectively. Assume for contradiction that firm f is not filled to capacity and $\mu_W(f) \neq \mu_F(f)$. Then there is a worker w who is matched to f in μ_W but not in μ_F. Since μ_W is worker-optimal, w prefers f to its match in μ_F. Since f is not filled to capacity, (w, f) forms a blocking pair of Type 2a with respect to μ_F, giving a contradiction. □

3 Incentive Compatibility

In this section we will study incentive compatibility properties of the Deferred Acceptance Algorithm for all three settings. Theorem 17 showed that if workers propose, then the matching computed is worker optimal in the sense that each worker is matched to the best firm in its realm of possibilities. However, for an individual worker this "best" firm may be very low in its preference list, see Exercise 6. If so, this worker may have incentive to cheat, i.e., manipulate its preference list in order to get a better match.

A surprising fact about the DA Algorithm is that this worker will not be able to get a better match by falsifying its preference list. Hence its best strategy is to report its true preference list. Moreover, this holds no matter what preference lists the rest of the workers report. Thus the worker-proposing DA Algorithm is *dominant-strategy incentive compatible (DSIC)* for workers.

This ground-breaking result on incentive compatibility opened up the DA Algorithm to a host of highly consequential applications. An example is its use for matching students to public schools in big cities, such as NYC and Boston, with hundreds of thousands of students seeking admission each year into hundreds of schools. Previously, Boston was using the Immediate Acceptance Algorithm which did not satisfy incentive compatibility. It therefore led to much guessing and gaming, making the process highly stressful for the students and their parents. With the use of the student-proposing DA Algorithm, each student is best off simply reporting her true preference list. For further details on this application, see Chapter **??**.

The proof of Theorem 32, showing DSIC for Setting I, is quite non-trivial and intricate, and more complexity is introduced in Setting II. In this context, the advantage of partitioning the problem into the three proposed settings should become evident.

The rest of the picture for incentive compatibility of the DA Algorithm is as follows. In Setting I, the worker-proposing DA Algorithm is not DSIC for firms; see Exercise 6. The picture is identical in Setting II and Setting III for the worker-proposing DA Algorithm. However, Setting III is asymmetrical for workers and firms, since firms have capacities. For this setting, Theorem 35 establishes that there is no mechanism that is DSIC for firms.

3.1 Proof of DSIC for Setting I

The following lemma will be critical to proving Theorem 32; it guarantees a blocking pair with respect to an arbitrary perfect matching μ. Observe that the blocking pair involves a worker which does not improve its match in going from μ_W to μ.

Lemma 29 (Blocking Lemma)**.** *Let μ_W be the worker-optimal stable matching under preferences $\succ$ and let μ be an arbitrary perfect matching, not necessarily stable. Further let W' be the set of workers who prefer their match under μ to their match under μ_W, i.e., $W' = \{w \in W \mid \mu(w) \succ_w \mu_W(w)\}$, and assume that $W' \neq \emptyset$. Then $W' \neq W$ and there exist $w \in (W \setminus W')$ and $f \in \mu(W')$ such that (w, f) is a blocking pair for μ.*

Proof. Clearly, for $w \in (W \setminus W')$, $\mu_W(w) \succeq_w \mu(w)$. Two cases arise naturally: whether the workers in W' get better matches in μ over μ_W by simply trading partners, i.e., whether $\mu(W') = \mu_W(W')$ or not. We will study the two cases separately.

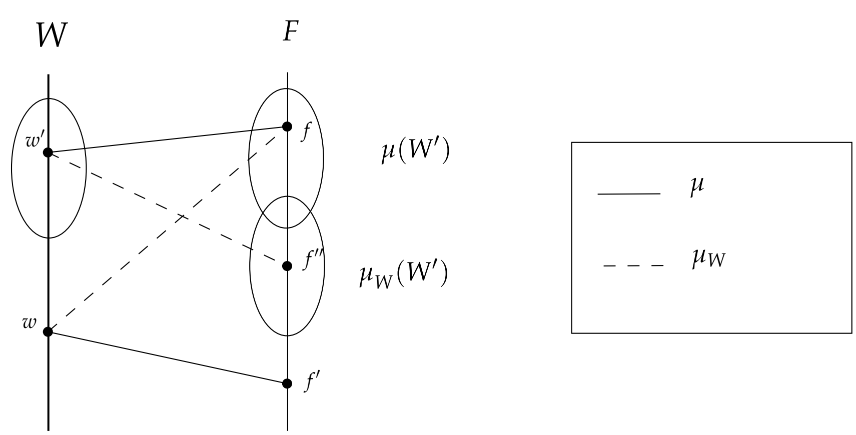


Figure 2

Case 1: $\mu(W') \neq \mu_W(W')$.

Since $|\mu(W')| = |\mu_W(W')| = |W'|$ and $(\mu(W') \setminus \mu_W(W')) \neq \emptyset$, therefore $W' \neq W$. Pick any $f \in (\mu(W') \setminus \mu_W(W'))$ and let $w = \mu_W(f)$. Now $w \in (W \setminus W')$ since if $f \notin \mu_W(W')$ then $\mu_W(f) \notin W'$. We will show that (w, f) is a blocking pair for μ. Towards this end, we will identify several other workers and firms.

Let $w' = \mu(f)$; since $f \in \mu(W')$, $w' \in W'$. Let $f'' = \mu_W(w')$; clearly, $f'' \in \mu_W(W')$. Finally let $f' = \mu(w)$; since $w \in (W \setminus W')$, $f' \in (F \setminus \mu(W'))$. Figure 2 will be helpful in visualizing the situation.

Finally we need to show that $f \succ_w f'$ and $w \succ_f w'$. The first assertion follows on observing that $w \in (W \setminus W')$ and $f \neq f'$. Assume that the second assertion is false, i.e., that $w' \succ_f w$. Now $f \succ_{w'} f''$, since $w' \in W'$, $f = \mu(w')$ and $f'' = \mu_W(w')$. But this implies that (w', f) is a blocking pair for stable matching μ_W. The contradiction proves that $w \succ_f w'$. Hence (w, f) is a blocking pair for μ.

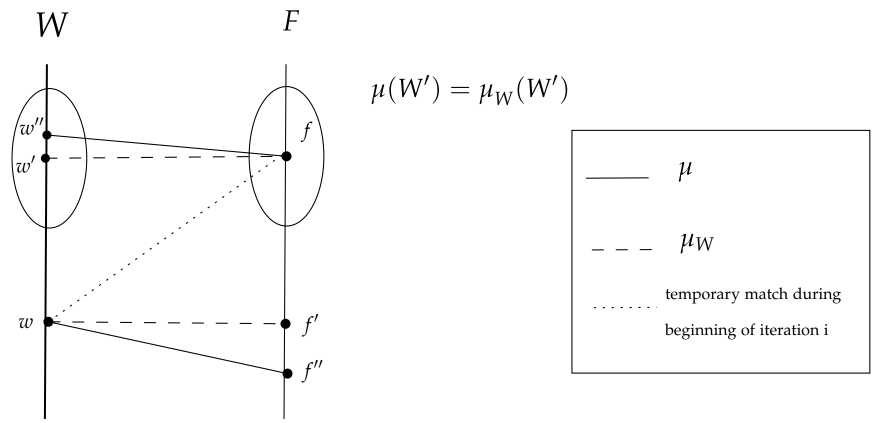


Figure 3

Case 2: $\mu(W') = \mu_W(W')$.

In this case, unlike the previous one, we will critically use the fact that μ_W is the matching produced by the DA Algorithm with workers proposing. Let i be the last iteration of the DA Algorithm in which a worker, say $w' \in W'$, first proposes to its eventual match; let the latter be $f \in \mu_W(W')$. Let $w'' = \mu(f)$. Since $f \in \mu(W')$, $w'' \in W'$.

By definition of W', $f = \mu(w'') \succ_{w''} \mu_W(w'')$. Therefore, w'' must have proposed to f before iteration i and subsequently moved on to its eventual match under μ_W no later than iteration i. Now, by Observation 5, f must keep getting proposals in each iteration after w'' proposed to it. In particular, assume that at the end of iteration $(i-1)$, f had tentatively accepted the proposal of worker w. In iteration i, f will reject w and w will propose to its eventual match in iteration $(i+1)$ or later. Since w' is the last worker in W' to propose to its eventual match, $w \notin W'$. Therefore $W' \neq W$. We will show that (w, f) is a blocking pair for μ. Figure 3 will be helpful in visualizing the situation.

Since f must have rejected w'' before tentatively accepting the proposal of w, $w \succ_f w''$. Let $f' = \mu_W(w)$ and $f'' = \mu(w)$. In the DA Algorithm, w had proposed to f before being finally matched to f', therefore $f \succ_w f'$. Since $w \in (W \setminus W')$, $f' \succeq_w f''$. Therefore $f \succ_w f''$. Together with the assertion $w \succ_f w''$, we get that (w, f) is a blocking pair for μ. □

Notation 30. Assume that a worker w reports a modified list; let us denote it by $\succ'_w$. Also assume that all other workers and all firms report their true preference lists. Define for all $x \in F \cup W$,

$$\succ'_x \; = \begin{cases} \succ_x & \text{if } x \in F \text{ or } x \in W \setminus \{w\}, \\ \succ'_w & \text{if } x = w \end{cases}$$

Let μ_W and μ'_W be the worker-optimal stable matchings under the preferences $\succ$ and $\succ'$, respectively. We will use these notions to state and prove Theorem 32. However, first we will give the following straightforward observation.

Observation 31. *Let $(W, F, \succ)$ be an instance of stable matching and consider alternative preference lists $\succ'_w$ for every worker $w \in W$ and $\succ'_f$ for every firm $f \in F$. Assume that we are given a perfect matching μ which has a blocking pair (w, f) with respect to the preferences $\succ'$. Moreover, assume that w and f satisfy: $\succ'_w = \succ_w$ and $\succ'_f = \succ_f$. Then (w, f) is a blocking pair in μ with respect to $\succ$ as well.*

Theorem 32. *Let $\succ$ and $\succ'$ be the preference lists defined above and let μ_W and μ'_W be the worker-optimal stable matchings under these preferences, respectively. Then*

$$\mu_W(w) \succeq_w \mu'_W(w),$$

i.e., the match of w, with respect to its original preference list, does not improve if it misrepresents its list as $\succ'_w$.

Proof. We will invoke Lemma 29; for this purpose, denote μ'_W by μ. Suppose w prefers its match in μ to its match in μ_W. Let $W' = \{w \in W \mid \mu(w) \succ_w \mu_W(w)\}$; clearly $w \in W'$ and therefore $W' \neq \emptyset$. Now by Lemma 29, there is a blocking pair (w', f) for μ with respect to the preferences $\succ$, with $w' \notin W'$; clearly, $w' \neq w$.

Since $\succ'$ and $\succ$ differ only for w, by Observation 31, (w', f) is a blocking pair for μ with respect to $\succ'$ as well. This contradicts the fact that μ is a stable matching with respect to $\succ'$. □

3.2 DSIC for Setting II

While studying incentive compatibility for the case of incomplete preference lists, we will allow a worker w to not only alter its preference list over its neighbors in graph G but to also alter its set of neighbors, i.e., to alter G itself. For this reason, it will be more convenient to define the preference list of each worker over the set $F \cup \{\perp\}$ and of each firm over the set $W \cup \{\perp\}$, as stated in the footnote to Definition 1 in Section 1.

Let $\succ$ denote the original preference lists of workers and firms and let μ_W denote the worker-optimal stable matching under preferences $\succ$. The definition of *blocking pair* with respect to a matching μ is changed in one respect only, namely $\mu(w) = \perp$ or $\mu(f) = \perp$ is allowed. Thus (w, f) is a blocking pair if and only if $(w, f) \notin \mu$, $\mu(w) \succ_w f$ and $\mu(f) \succ_f w$.

The only change needed to the statement of the Blocking Lemma, stated as Lemma 29 for Setting I, is that μ is an arbitrary matching, i.e., not necessarily perfect. Once again, the proof involves the same two cases presented in Lemma 29. The proof of the first case changes substantially and is given below.

Proof. **Case 1:** $\mu(W') \neq \mu_W(W')$.

For $w \in W'$, $\mu(w) \succ_w \mu_W(w)$, therefore $\mu(w) \neq \perp$. However, $\mu_W(w) = \perp$ is allowed. Therefore $|\mu(W')| = |W'| \geq |\mu_W(W')|$. Hence, $\mu(W') \not\subset \mu_W(W')$, and since $\mu(W') \neq \mu_W(W')$, we get that $(\mu(W') \setminus \mu_W(W')) \neq \emptyset$. Pick any $f \in (\mu(W') \setminus \mu_W(W'))$. Clearly $\mu(f) \neq \perp$; let $w' = \mu(f)$. Let $\mu_W(w') = f''$, where $f'' = \perp$ is possible. By definition of W', $f \succ_{w'} f'' = \mu_W(w')$.

Assume for contradiction that $\mu_W(f) = \perp$. Since $\mu(f) = w'$, $w' \succ_f \perp$. Therefore we get that $w' \succ_f \mu_W(f)$. If so, (w', f) is a blocking pair for μ_W, leading to a contradiction. Therefore $\mu_W(f) \neq \perp$. Let $\mu_W(f) = w$. Clearly, $w \notin W'$. Hence $W' \neq W$. Let $\mu(w) = f'$, where $f' = \perp$ is possible. Clearly, $f \neq f'$ and $f \succ_w f'$. This together with the assertion $w \succ_f w'$ gives us that (w, f) is a blocking pair for μ. □

Consider Case 2, namely $\mu(W') = \mu_W(W')$. Since for each $w \in W'$, $\mu(w) \succ_w \mu_W(w)$, w cannot be unmatched under μ. Therefore, in this case, $\forall w \in W'$, w is matched in both μ and μ_W. The rest of proof is identical to that in Lemma 29, other than the fact that $f' = \perp$ and $f'' = \perp$ are allowed. We will not repeat the proof here. This establishes the Blocking Lemma for Setting II.

As in Setting I presented in Section 3.1, we will adopt the notation $\succ$ and $\succ'$ given in Notation 30. Let μ'_W be the worker-optimal stable matchings under the preferences $\succ'$. Note that Observation 6 still holds. Finally, the statement of Theorem 32 carries over without change to this setting and its proof is also identical, provided one uses the slightly-modified statement of Blocking Lemma.

3.3 DSIC for Setting III

For Setting III, DSIC for workers follows easily using the reduction from Setting II to Setting III, given in Section 2.3, and the fact that the worker-proposing DA Algorithm is DSIC for workers. However, this setting is asymmetric for workers and firms, since firms have capacities. Therefore, we need to study incentive compatibility of the firm-proposing DA Algorithm as well. The answer is quite surprising: not only is this algorithm not DSIC, but in fact no mechanism can be DSIC for this setting.

Example 33. Let the set of workers be $W = \{w_1, w_2, w_3, w_4\}$ and firms be $F = \{f_1, f_2, f_3\}$. Firm f_1 has capacity 2 and f_2 and f_3 have unit capacities. Preferences are given by:

$f_1 : w_1 \succ w_2 \succ w_3 \succ w_4$,
$f_2 : w_1 \succ w_2 \succ w_3 \succ w_4$,
$f_3 : w_3 \succ w_1 \succ w_2 \succ w_4$,
$w_1 : f_3 \succ f_1 \succ f_2$,
$w_2 : f_2 \succ f_1 \succ f_3$,
$w_3 : f_1 \succ f_3 \succ f_2$,
$w_4 : f_1 \succ f_2 \succ f_3$.

Consider the instance given in Example 33. If firms propose, the resulting stable matching, μ_F, assigns $\mu_F(f_1) = \{w_3, w_4\}$, $\mu_F(f_2) = \{w_2\}$ and $\mu_F(f_3) = \{w_1\}$.

Next consider matching μ with $\mu(f_1) = \{w_2, w_4\}$, $\mu(f_2) = \{w_1\}$ and $\mu(f_3) = \{w_3\}$. This matching is strictly preferred by all firms but it is not stable since (w_1, f_1) and (w_3, f_1) are both blocking pairs. However, if f_1 misrepresents its preferences as

$$f_1 : w_2 \succ' w_4 \succ' w_1 \succ' w_3,$$

then μ becomes stable and firm-optimal. Hence we get:

Lemma 34. *For the instance given in Example 33, there exists a matching μ which is not stable and which all firms strictly prefer to the firm-optimal stable matching μ_F. Moreover, there is a way for one firm to misrepresent its preferences so that μ becomes stable.*

The next theorem follows.

Theorem 35. *For the stable matching problem in Setting III with capacitated firms, there is no mechanism that is DSIC for firms.*

4 The Lattice of Stable Matchings

The notions of worker-optimal and worker-pessimal stable matchings, defined in Section 2.1, indicate that the set of all stable matchings of an instance has structure. It turns out that these notions form only the tip of the iceberg! Below we will define the notion of a *finite distributive lattice* and will prove that the set of stable matchings of an instance forms such a lattice. Together with the non-trivial notion of *rotation* and Birkhoff's Representation Theorem, this leads to an extremely rich collection of structural properties and efficient algorithms which find use in important applications.

4.1 The lattice for Setting I

Definition 36. Let S be a finite set, $\geq$ be a reflexive, anti-symmetric, transitive relation on S and $\pi = (S, \geq)$ be the corresponding partially ordered set. For any two elements $a, b \in S$, $u \in S$ is said to be an *upper bound* of a and b if $u \geq a$ and $u \geq b$. Further, u is said to be a *least upper bound* of a and b if $u' \geq u$ for any upper bound u' of a and b. The notions of *(greatest) lower bound* of two elements are analogous. Partial order π is said to be a *lattice* if any two elements $a, b \in S$ have a unique least upper bound and a unique greatest lower bound. If so, these will be called the *join* and *meet* of a and b and will be denoted by $a \vee b$ and $a \wedge b$, respectively, and the partial order will typically be denoted by $\mathcal{L}$. Finally, $\mathcal{L}$ is said to be a *finite distributive lattice*, abbreviated *FDL*, if for any three elements $a, b, c \in S$, the distributive property holds, i.e.,

$$a \vee (b \wedge c) \;=\; (a \vee b) \wedge (a \vee c) \quad \text{and} \quad a \wedge (b \vee c) \;=\; (a \wedge b) \vee (a \wedge c).$$

Birkhoff's Representation Theorem, mentioned above, holds for FDLs. We next define a natural partial order on the set of stable matchings of an instance and show that it forms such a lattice.

Definition 37. Let S_μ be the set of stable matchings of a given instance in Setting I. Define a relation $\geq$ on S_μ as follows: for $\mu, \mu' \in S_\mu$, $\mu \geq \mu'$ if and only if every worker w weakly prefers her match in μ to her match in μ', i.e.,

$$\forall w \in W : \; \mu(w) \succeq_w \mu'(w).$$

Theorem 41 shows that the partial order $\mathcal{L}_\mu = (S_\mu, \geq)$ is a finite distributive lattice. $\mathcal{L}_\mu$ will be called the *stable matching lattice* for the given instance.

Let μ and μ' be two stable matchings. We define the following four operations: For worker w, $\max\{\mu(w), \mu'(w)\}$ is the firm which w weakly prefers among $\mu(w)$ and $\mu'(w)$, and $\min\{\mu(w), \mu'(w)\}$ is the firm that w weakly dislikes, where "dislikes" is the opposite of the relation prefers. For a firm f, $\max\{\mu(f), \mu'(f)\}$ and $\min\{\mu(f), \mu'(f)\}$ are analogously defined.

Define two maps $M_W : W \to F$ and $M_F : F \to W$ as follows:

$$\forall w \in W: \; M_W(w) = \max\{\mu(w), \mu'(w)\} \quad \text{and} \quad \forall f \in F: \; M_F(f) = \min\{\mu(f), \mu'(f)\}.$$

Lemma 38. *$\forall w \in W$, if $M_W(w) = f$, then $M_F(f) = w$.*

Proof. Assume $\mu(w) \neq \mu'(w)$, since otherwise the proof is obvious. Let $\mu(w) = f$ and $\mu'(w) = f'$, and without loss of generality assume that $f \succeq_w f'$. Let $\mu'(f) = w'$; clearly $w \neq w'$. Now if $w \succeq_f w'$ then (w, f) is a blocking pair for μ', leading to contradiction. Therefore, $w' \succeq_f w$ and hence $M_F(f) = w$; see Figure 4 for an illustration. □

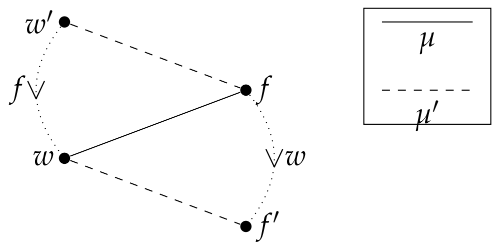


Figure 4: Figure for proof of Lemma 38.

Corollary 39. *M_W and M_F are both bijections, and $M_W = M_F^{-1}$.*

As a consequence of Corollary 39, M_W is a perfect matching on $W \cup F$; denote it by μ_1. Analogously, mapping each worker w to $\min\{\mu(w), \mu'(w)\}$ gives another perfect matching; denote it by μ_2. Observe that μ_2 matches each firm f to $\max\{\mu(f), \mu'(f)\}$, see Figure 5.

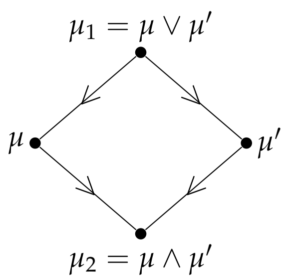


Figure 5: The meet and join of μ and μ'.

Lemma 40. *Matchings μ_1 and μ_2 are both stable.*

Proof. Assume that (w, f) is a blocking pair for μ_1. Let $\mu(w) = f'$ and $\mu'(w) = f''$ and assume without loss of generality that $f' \succeq_w f''$. Then $\mu_1(w) = \max\{\mu(w), \mu'(w)\} = f'$.

Let $\mu(f) = w'$ and $\mu'(f) = w''$. By the definition of map M_F, $w'' \succeq_f w'$ and $\mu_1(f) = w'$; observe that $w \neq w'$. Since (w, f) is a blocking pair for μ_1, $w \succeq_f w'$. This implies that (w, f) is a blocking pair for μ, leading to a contradiction. Hence μ_1 is stable. An analogous argument shows that μ_2 is also stable; see Figure 6 for an illustration. □

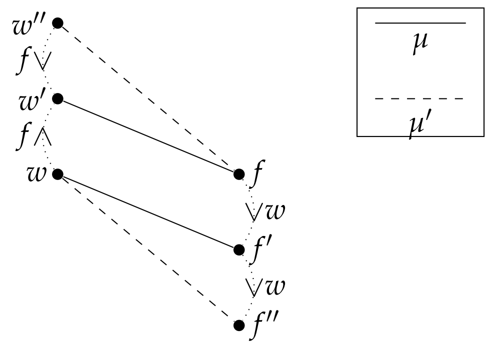


Figure 6: Figure for proof of Lemma 40.

Consider the partial order $\mathcal{L}_\mu = (S_\mu, \geq)$ defined in Definition 37. Clearly, μ_1 and μ_2 are an upper bound and a lower bound of μ and μ', respectively. It is easy to see that they are also the unique lowest upper bound and the unique greatest lower bound of μ and μ'. Therefore $\mathcal{L}_\mu$ supports the operations of meet and join given by:

$$\mu \vee \mu' = \mu_1 \quad \text{and} \quad \mu \wedge \mu' = \mu_2,$$

Finally, it is easy to show that the operations of meet and join satisfy the distributive property, see Exercise 7. Hence we get:

Theorem 41. *The partial order $\mathcal{L}_\mu = (S_\mu, \geq)$ is a finite distributive lattice.*

Remark 42. If $\mu, \mu' \in S_\mu$ with $\mu > \mu'$, then by definition, workers get weakly better matches in μ as compared to μ'. The discussion presented above implies that firms get weakly worse matches.

Using the finiteness of $\mathcal{L}_\mu$, it is easy to show that there are two special matchings, say $\mu_\top, \mu_\bot \in S_\mu$, which we will call *top* and *bottom*, respectively, such that $\forall \mu \in S_\mu$, $\mu_\top \geq \mu$ and $\mu \geq \mu_\bot$. These stable matchings were already singled out in Section 2.1: $\mu_\top$ is the worker-optimal and firm-pessimal matching and $\mu_\bot$ is the firm-optimal and worker-pessimal matching, see Figure 7.

4.1.1 Rotations, and their use for traversing the lattice

Several applications require stable matchings which are not as "extreme" as $\mu_\top$ and $\mu_\bot$, i.e., they treat the two sets W and F more equitably. These can be found in the rest of the lattice. In this

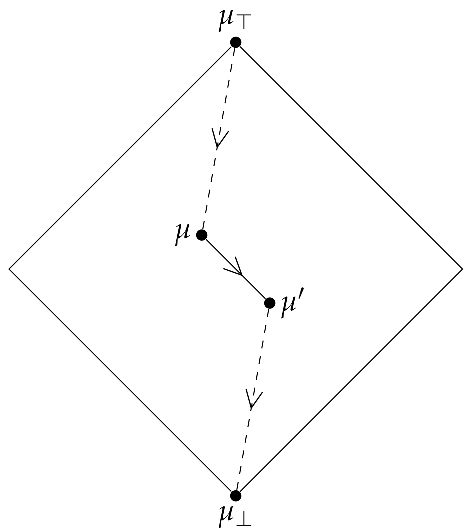


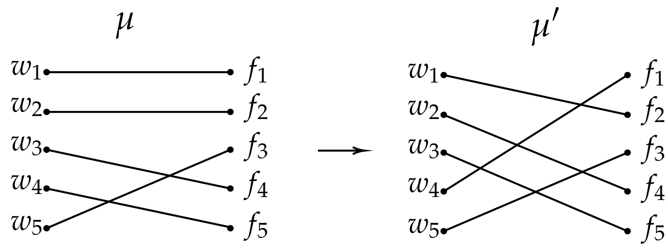
Figure 7: A path from the worker optimal matching, $\mu_\top$, to the firm optimal matching, $\mu_\bot$, in lattice $\mathcal{L}_\mu$. The edge (μ, μ') indicates that $\mu' = \rho(\mu)$, for a rotation ρ with respect to μ.

section we will define the notion of a rotation, which will help traverse the lattice. In particular, we will prove that rotations help traverse paths from $\mu_\top$ to $\mu_\bot$, as illustrated in Figure 7, with intermediate "vertices" on such a path being stable matchings and an "edge" (μ, μ') indicating that $\mu > \mu'$; see Lemma 45 and Corollary 46. By Remark 42, the intermediate matchings on any such path will gradually become better for firms and worse for workers. For an example of the use of rotations, see Exercise 12, which develops an efficient algorithm for finding a stable matching that treats workers and firms more equitably.

Definition 43. Fix a stable matching $\mu \neq \mu_\bot$ and define the function next : $W \to F \cup \{\boxtimes\}$ as follows: for worker w, find its most preferred firm, say f, such that f prefers w to $\mu(f)$. If such a firm exists, then $\text{next}(w) = f$ and otherwise $\text{next}(w) = \boxtimes$. A *rotation* ρ with respect to μ is an ordered sequence of pairs $\rho = \{(w_0, f_0), (w_1, f_1), \ldots (w_{r-1}, f_{r-1})\}$ such that for $0 \leq i \leq r-1$:

- $(w_i, f_i) \in \mu$, and
- $\text{next}(w_i) = f_{(i+1) \pmod r}$

By *applying rotation ρ to μ* we mean switching the matching of $w_0, \ldots, w_{r-2}, w_{r-1}$ to $f_1, \ldots, f_{r-1}, f_0$, respectively, and leaving the rest of μ unchanged. Clearly this results in a perfect matching; let us denote it by μ'. We will denote this operation as $\mu' = \rho(\mu)$. For example in the figure given below, rotation $\{(1,1), (2,2), (3,4), (4,5)\}$ applied to μ yields μ'. Observe that the matched edge $(5,3)$ is not in the rotation and remains unchanged in going from μ to μ'.



Lemma 44. *Let ρ be a rotation with respect to μ and let $\mu' = \rho(\mu)$. Then:*

1. *Workers get weakly worse matches and firms get weakly better matches in going from μ to μ'.*
2. *μ' is a stable matching.*
3. *$\mu > \rho(\mu)$.*

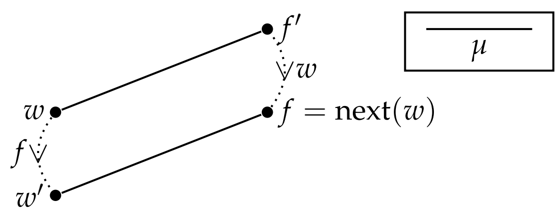


Figure 8

Proof. **1).** Suppose $\text{next}(w) = f$; clearly $\mu(f) \neq w$. Let $\mu(w) = f'$ and $\mu(f) = w'$. By the definition of next, $w \succ_f w'$. Observe that $f \succ_w f'$ is not possible, since then (w, f) will be a blocking pair for μ. Therefore, $f' \succ_w f$. Hence, after the rotation, the matching of f improved and that of w became worse. Clearly, this holds for all workers and firms in the rotation, see Figure 8.

2). Assume that μ' has a blocking pair, namely (w, f), where $\mu'(w) = f'$ and $\mu'(f) = w'$. From this blocking pair, we can infer that $w \succ_f w'$ and $f \succ_w f'$. Now there are two cases:

Case 1: $\mu(w) = f'$.
Assume that $\mu(f) = w''$, by part (1) of this lemma, we have $w' \succ_f w''$. The above-stated assertions give us $w \succ_f w''$, hence showing that (w, f) is a blocking pair for μ and leading to a contradiction, see Figure 9.

Case 2: $\mu(w) \neq f'$.
If $\mu(f) \neq w'$, then let $\mu(f) = w''$. Since f improves its match after the rotation, $w' \succeq_f w''$ and using the above-stated assertion, $w \succeq_f w''$. Therefore, whether or not $\mu(f) = w'$, we have that $w \succ \mu(f)$.

Clearly, $\text{next}(w) = f'$. Since f prefers w to its match in μ and $f \neq \text{next}(w)$, we get that $f' \succ_w f$, hence contradicting the above-stated assertion that $f \succ_w f'$, see Figure 10.

3). This follows from the previous two statements. □

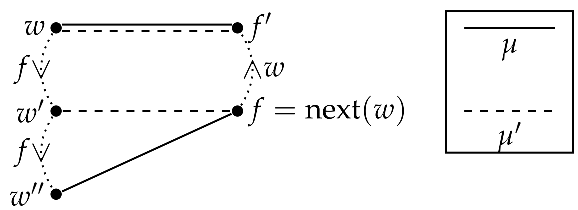


Figure 9: Figure for proof of Case 1 in Lemma 44.

Assume that next(w) $\neq \perp$, i.e., next(w) is a firm. We note that this does not guarantee that w will be in a rotation, see Exercise 9. For that to happen, the "cycle must close", i.e., for some r, $\text{next}(w_{r-1}) = f_0$.

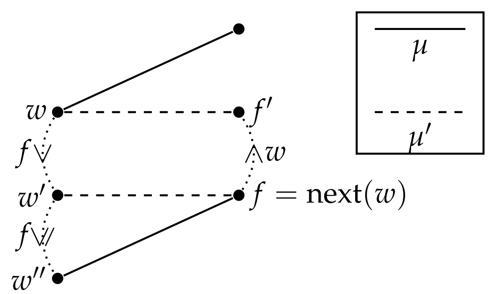


Figure 10: Figure for proof of Case 2 in Lemma 44.

The next lemma will justify Figure 7 via Corollary 46.

Lemma 45. *Let $\mu > \mu'$. Then there exists a rotation ρ with respect to μ such that $\mu > \rho(\mu) \geq \mu'$.*

Proof. Define map $g : W \to W \cup \{\boxtimes\}$ as follows:

$$g(w) = \begin{cases} \mu(\text{next}(w)) & \text{if next}(w) \in F \\ \boxtimes & \text{otherwise.} \end{cases}$$

Let $W' = \{w \in W \mid \mu(w) \succ \mu'(w)\}$. We will prove that the range of g when restricted to W' is W' and for $w \in W'$, $g(w) \neq w$.

Let $w \in W'$. We will first prove that next$(w) \in F$, hence showing that $g(w) \in W$. Let $\mu'(w) = f'$ and $\mu(f') = w''$. Since $\mu > \mu'$, μ is weakly better than μ' for workers. Therefore, $f' \succ_{w''} \mu'(w'')$. If $w'' \succ_{f'} w$, then (w'', f') will be a blocking pair for μ', leading to a contradiction. Therefore, $w \succ_{f'} w''$. Hence there is a firm that likes w better than its own match under μ. Among such firms, let f be one that w prefers most. Then next$(w) = f$ and $g(w) = \mu(f) = w'$, say.

Next, we will show that $\mu(w') \succ \mu'(w')$, hence getting that $w' \in W'$; see Figure 11. Suppose not. Since $\mu > \mu'$ and μ is not strictly better than μ' for w', we must have that $\mu'(w') = \mu(w') = f$. Let $\mu'(w) = f'$. Since $\mu > \mu'$, we have that $w \succ_{f'} \mu(f')$. Therefore f' is a firm that prefers w

to its match under μ. However, since f is the most preferred such firm per w, we get $f \succ_w f'$. Furthermore, since $\mu(f) = w'$, $w \succ_f w'$. Therefore (w, f) is a blocking pair for μ', leading to a contradiction. Therefore, $\mu(w') \succ \mu'(w')$ and hence $w' \in W'$. Clearly w' is distinct from w, hence giving $g(w) \neq w$.

Finally, we will use the map $g : W' \to W'$ to complete the proof. Start with any worker $w \in W'$ and repeatedly apply g until a worker is encountered a second time. This gives us a "cycle", i.e., a sequence of workers $w_0, w_1, \ldots, w_{r-1}$ such that for $0 \leq i \leq r-1$, $g(w_i) = w_{i+1(\bmod r)}$. Then $\rho = \{(w_0, f_0), (w_1, f_1), \ldots (w_{r-1}, f_{r-1})\}$ is a rotation with respect to μ, where $f_i = \mu(w_i)$ for $0 \leq i \leq r-1$. Clearly, $\mu > \rho(\mu) \geq \mu'$. □

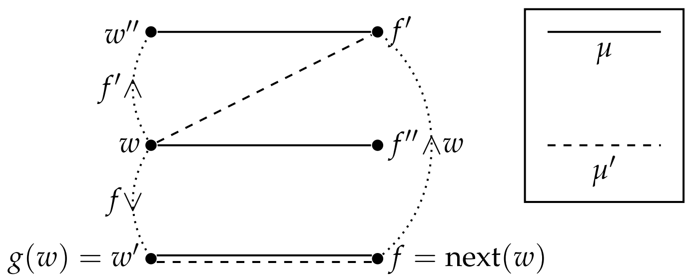


Figure 11: Figure to illustrate why $w' \in W'$ in the proof of Lemma 45.

Corollary 46. *The following hold:*

1. *Let μ be a stable matching such that $\mu \neq \mu_\perp$. Then there is a rotation ρ with respect to μ.*
2. *Start with $\mu_\top$ as the "current matching" and successively apply a rotation with respect to the current matching. This process will terminate at $\mu_\perp$.*

Let $G_\mu = (S_\mu, E_\mu)$ be a directed graph with vertex set S_μ and $(\mu, \mu') \in E_\mu$ if and only if there is a rotation ρ with respect to μ such that $\rho(\mu) = \mu'$. Then any path from $\mu_\top$ to $\mu_\perp$ is obtained by the process given in Corollary 46.

Definition 47. Let ρ be a rotation with respect to stable matching μ and let $\rho(\mu) = \mu'$. Then the inverse of the map ρ is denoted by ρ^{-1}. We will call ρ^{-1} the *inverse rotation* with respect to stable matching μ'. Clearly, $\rho^{-1}(\mu') = \mu$.

Clearly inverse rotations help traverse paths from $\mu_\perp$ to $\mu_\top$ in G_μ, and a combination of rotations and inverse rotations suffice to find a path from any matching to any other matching in G_μ.

4.1.2 Rotations correspond to join-irreducibles

Definition 48. A stable matching μ is said to be *join-irreducible* if $\mu \neq \mu_\perp$ and μ is not the join of any two stable matchings. Let μ and μ' be two stable matchings. We will say that μ' is the *direct successor* of μ if $\mu > \mu'$ and there is no stable matching μ'' such that $\mu > \mu'' > \mu'$; if so, we will denote this by $\mu \rhd \mu'$.

Let μ be a join-irreducible stable matching. By Lemma 44 and Corollary 46 there is a unique rotation ρ with respect to μ. Let $\rho(\mu) = \mu'$. Clearly μ' is the unique stable matching such that $\mu \triangleright \mu'$.

Lemma 49. *Let ρ_1 and ρ_2 be two distinct rotations with respect to stable matching μ, and let $\mu_1 = \rho_1(\mu)$ and $\mu_2 = \rho_2(\mu)$. Then the following hold:*

1. *ρ_1 and ρ_2 cannot contain the same worker-firm pair.*
2. *ρ_1 and ρ_2 are rotations with respect to μ_2 and μ_1, respectively.*

Proof. **1).** Clearly, $\mu > \mu_1$ and $\mu > \mu_2$. Consider the set W' and the map $g : W' \to W'$ defined in the proof of Lemma 45. It is easy to see that ρ_1 and ρ_2 correspond to two disjoint "cycles" in this map, hence giving the lemma.

2). Since ρ_1 and ρ_2 correspond to disjoint "cycles", applying one results in a matching to which the other can be applied. □

Notation 50. We will denote the set of all rotations used in lattice $\mathcal{L}_\mu$ by $\mathcal{R}_\mu$, and the set of join-irreducibles of $\mathcal{L}_\mu$ by $\mathcal{J}_\mu$.

Lemma 51. *Let $\rho \in \mathcal{R}_\mu$. Then there is a join-irreducible, say μ, such that ρ is the unique rotation with respect to μ.*

Proof. Let $S_\rho = \{\nu \in S_\mu \mid \rho \text{ is a rotation with respect to } \nu\}$. Let μ be the meet of all matchings in S_ρ. Clearly $\mu \in S_\rho$ and hence ρ is a rotation with respect to μ. Suppose μ is not join-irreducible and let ρ' be another rotation with respect to μ. By the second part of Lemma 49, ρ is a rotation with respect to $\rho'(\mu)$, therefore $\rho'(\mu) \in S_\rho$. This implies that $\rho'(\mu) \geq \mu$. But $\mu > \rho'(\mu)$, giving a contradiction. Uniqueness follows from the fact that μ is join-irreducible. □

Lemma 52. *Two rotations ρ and ρ' cannot contain the same worker-firm pair.*

Proof. Suppose rotations ρ and ρ' both contain the same worker-firm pair, say (w, f). By Lemma 51, there are two join-irreducible stable matchings μ and ν such that ρ and ρ' are rotations with respect to these matchings. Let $\mu \triangleright \mu'$ and $\nu \triangleright \nu'$.

Consider the matching $\mu \vee \nu$. Since $(w, f) \in \mu$ and $(w, f) \in \nu$, $(w, f) \in (\mu \vee \nu)$. On the other hand, since (w, f) is in rotations ρ and ρ', w is matched to a better firm than f in μ' and in ν'. Therefore w is matched to a better firm than f in $\mu' \vee \nu'$. But since μ and ν are join irreducible, $(\mu' \vee \nu') = (\mu \vee \nu)$, leading to a contradiction. Hence ρ and ρ' cannot contain the same worker-firm pair. □

Lemmas 51 and 52, together with the facts that there are n^2 worker-firm pairs and each rotation has at least two worker-firm pairs, give:

Corollary 53. *The following hold:*

1. *There is a bijection $f : \mathcal{J}_\mu \to \mathcal{R}_\mu$ such that if $f(\mu) = \rho$ then ρ is the unique rotation with respect to μ.*

2. $|\mathcal{R}_\mu| \leq n^2/2$.

4.1.3 Birkhoff's Representation Theorem

FDLs arise in diverse settings; Definition 54 gives perhaps the simplest of these. A consequence of Birkhoff's Representation Theorem is that each FDL is isomorphic to a canonical FDL. In this section we will prove this for stable matching lattices; the general statement follows along similar lines.

Definition 54. Let S be a finite set and $\mathcal{F}$ be a family of subsets of S which is closed under union and intersection. Denote the partial order $(\mathcal{F}, \supseteq)$ by $\mathcal{L}_\mathcal{F}$. Then $\mathcal{L}_\mathcal{F}$ is a FDL with meet and join being given by:

$$A \wedge B = A \cap B \quad \text{and} \quad A \vee B = A \cup B,$$

for any two sets $A, B \in \mathcal{F}$. $\mathcal{L}_\mathcal{F}$ will be called a *canonical finite distributive lattice*.

Definition 55. The projection of $\mathcal{L}_\mu$ onto $\mathcal{J}_\mu$ will be called a *join-irreducible partial order* and will be denoted by $\pi_\mu = (\mathcal{J}_\mu, \geq)$. We will say that $S \subseteq \mathcal{J}_\mu$ is a *lower set of* π_μ if it satisfies: if $\mu \in S$ and $\mu > \mu'$ then $\mu' \in S$.

Let $\mathcal{F}_\pi$ be the family of subsets of $\mathcal{J}_\mu$ consisting of all lower sets of π_μ. It is easy to see that $\mathcal{F}_\pi$ is closed under union and intersection, and therefore $\mathcal{L}_\pi = (\mathcal{F}_\mu, \supseteq)$ is a canonical FDL.

Theorem 56. *The lattice $\mathcal{L}_\mu$ is isomorphic to $\mathcal{L}_\pi$, i.e., there is a bijection $f_\mu : S_\mu \rightarrow \mathcal{F}_\pi$ such that $\mu \succeq \mu'$ if and only if $f_\mu(\mu) \supseteq f_\mu(\mu')$.*

Proof. Define function $f_\mu : S_\mu \rightarrow \mathcal{F}_\pi$ as follows: for $\mu \in S_\mu$, $f(\mu)$ is the set of all join-irreducibles ν such that $\mu \succ \nu$; let this set be S. Then S is a lower set of π_μ since if $\mu_1, \mu_2 \in \mathcal{J}_\mu$, with $\mu_1 \in S$ and $\mu_1 \succ \mu_2$ then $\mu \succ \mu_2$, therefore giving that $\mu_2 \in S$. Hence $S \in \mathcal{F}_\pi$. Next define function $g : \mathcal{F}_\pi \rightarrow S_\mu$ as follows: for a lower set S of π_μ, $g(S)$ is the join of all join-irreducibles $\nu \in S$.

We first show that the compositions $g \bullet f$ and $f \bullet g$ are both the identity function, thereby showing that f and g are both bijections. Then $f_\mu = f$ is the required bijection.

Let $\mu \in S_\mu$ and let $f(\mu) = S$. For the first composition, we need to show that $g(S) = \mu$. There exist $j_1, \ldots, j_k$, join-irreducibles of $\mathcal{L}_\mu$, such that $\mu = (j_1 \vee \ldots \vee j_k)$[5]. Clearly, $j_1, \ldots, j_k \in S$ and therefore $g(S) \succeq (j_1 \vee \ldots \vee j_k)$. Furthermore, $\mu \succeq g(S)$. Therefore, $\mu \succeq g(S) \succeq (j_1 \vee \ldots \vee j_k) = \mu$. Hence $g(S) = \mu$.

Let S be a lower set of π_μ, $j_1, \ldots, j_k$ be the join-irreducibles in S and μ be $g(S)$, i.e., the join of these join-irreducibles. Let j be a join-irreducible of $\mathcal{L}_\mu$ such that $\mu \succeq j$. For the second composition, we need to show that $j \in S$, since then $f(\mu) = S$. Now,

$$j \wedge \mu \;=\; j \wedge (j_1 \vee \ldots \vee j_k) \;=\; (j \wedge j_1) \vee \ldots \vee (j \wedge j_k),$$

where the second equality follows from the distributive property. Since j is a join-irreducible, it cannot be the join of two or more elements. Therefore, $j = j \wedge j_i$, for some i. But then $j_i \succeq j$, therefore giving $j \in S$.

[5]Observe that every $\mu \in S_\mu$ can be written as the join of a set of join-irreducibles as follows: If μ is a join-irreducible, there is nothing to prove. Otherwise, it is the join of two or more matchings. Continue the process on each of those matchings.

Finally, the definitions of f and g give that $\mu \succeq \mu'$ if and only if $f_\mu(\mu) \supseteq f_\mu(\mu')$. $\square$

Observe that an instance may have exponentially many, in n, stable matchings, hence leading to an exponentially large lattice. On the other hand, by Corollary 53, π_μ, which encodes this lattice, has a polynomial sized description. The precise way that π_μ encodes $\mathcal{L}_\mu$ is clarified in Exercise 10.

Corollary 57. *There is a succinct description of the stable matching lattice.*

4.2 The lattice for Settings II and III

The entire development on the lattice of stable matchings for Setting I, presented in Section 4.1, can be easily ported to Setting II with the help of the Rural Hospital Theorem, Theorem 23, which proves that the set of workers and firms matched is the same in all stable matchings.

Let these sets be $W' \subseteq W$ and $F' \subseteq F$. Let $w \in W'$. Clearly, it suffices to restrict the preference list of w to F' only. If this list is not complete over F', simply add the missing firms of F' at the end; since w is never matched to these firms, their order does not matter. Applying this process to each worker and each firm yields an instance of stable matching over W' and F' which is in Setting I. The lattice for this instance is also the lattice for the original instance in Setting II.

The lattice for Setting III requires the reduction from Setting III to Setting II, given in Section 2.3, and the Rural Hospital Theorem for Setting III, given in Theorem 28. The latter proves that if a firm is not matched to capacity in a stable matching, then it is matched to the same set of workers in all stable matchings. However, a firm that is matched to capacity may be matched to different sets of workers in different stable matchings.

Assume that firm f is matched to capacity in μ_1 and μ_2. A new question that arises is: which of these two matchings does f prefer? Observe that questions of this sort have natural answers for workers and firms in the previous settings (and for workers in Setting III) and these answers were a key to formulating the lattice structure. The new difficulty is the following: if the two sets $\mu_1(f)$ and $\mu_2(f)$ are interleaved in complicated ways, when viewed with respect to the preference order of f, we will have no grounds for declaring one matching better than the other. Lemma 60 shows that if μ_1 and μ_2 are stable matchings, then such complications do not arise.

Definition 58. Fix a firm f. For $W' \subseteq W$, define $\min(W')$ to be the worker whom f prefers the least among workers in W'. For $W_1 \subseteq W$ and $W_2 \subseteq W$, we will say that f *prefers* W_1 to W_2 if f prefers every worker in W_1 to every worker in W_2 and we will denote this as $W_1 \gg_f W_2$.

$$W_1 = \{w \in W \mid \mu_1(w) \succ_w \mu_2(w)\} \quad \text{and} \quad W_2 = \{w \in W \mid \mu_2(w) \succ_w \mu_1(w)\}.$$

Also let

$$F_1 = \{f \in F \mid \min(\mu_2(f) - \mu_1(f)) \gg_f \min(\mu_1(f) - \mu_2(f))\} \quad \text{and}$$
$$F_2 = \{f \in F \mid \min(\mu_1(f) - \mu_2(f)) \gg_f \min(\mu_2(f) - \mu_1(f))\}.$$

Lemma 59. *Let $(w, f) \in (\mu_1 - \mu_2)$. Then,*

1. $w \in W_1 \implies f \in F_1$

2. $w \in W_2 \implies f \in F_2$.

Lemma 60. *Let μ_1 and μ_2 be two stable matchings and f be a firm that is matched to capacity in both matchings. Then one of these possibilities must hold:*

1. $\mu_1(f) - \mu_2(f) \gg_f \mu_2(f)$ *or*
2. $\mu_2(f) - \mu_1(f) \gg_f \mu_1(f)$.

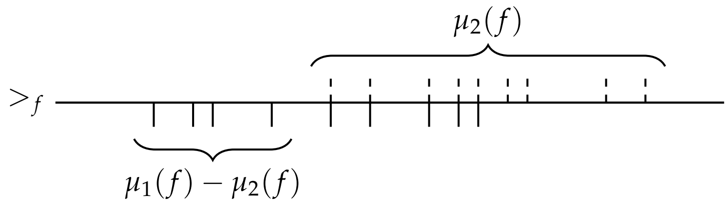


Figure 12: Figure illustrating the first possibility in Lemma 60. The horizontal line indicates the preference list of f, in decreasing order from left to right. The markings below the line indicate workers in set $\mu_1(f)$ and those above indicate workers in $\mu_2(f)$.

Definition 61. Let μ_1 and μ_2 be two stable matchings and f be a firm that is matched to capacity in both matchings. Then f prefers μ_1 to μ_2 if the first possibility in Lemma 60 holds and it prefers μ_2 to μ_1 otherwise. We will denote these as $\mu_1 \succ_f \mu_2$ and $\mu_2 \succ_f \mu_1$, respectively.

Using Definition 61, whose validity is based on Lemma 60, we get a partial order on the set of stable matchings for Setting III. Using the reduction stated above and facts from Section 4.1, this partially ordered set forms a FLD.

5 Linear Programming Formulation

The stable matching problem in Setting I admits a linear programming formulation in which the polyhedron defined by the constraints has integral optimal vertices, i.e., the vertices of this polyhedron are stable matchings. This yields an alternative way of computing a stable matching in polynomial time using well-known ways of solving LPs. LPs for Settings II and III follow using the Rural Hospital Theorem 23 and the reduction from Setting III to Setting II given in Section 2.3, respectively.

5.1 LP for Setting I

A sufficient condition for a worker-firm pair, (w, f) to not form a blocking pair with respect to a matching μ is that if w is matched to firm f' such that $f \succ_w f'$ then f should be matched to a worker w' such that $w' \succ_f w$. The fractional version of this condition appears in the third constraint of LP (1); see also Exercise 14. The first two constraints ensure that each worker and firm is fully matched. Observe that the objective function is simply 0.

$$
\begin{aligned}
\max \quad & 0 \\
\text{s.t.} \quad & \sum_{w} x_{wf} = 1 \quad \forall w \in W, \\
& \sum_{f} x_{wf} = 1 \quad \forall f \in F, \\
& \sum_{f \succ_w f'} x_{wf'} - \sum_{w' \succ_f w} x_{w'f} \leq 0 \quad \forall w \in W,\ \forall f \in F, \\
& x_{wf} \geq 0 \quad \forall w \in W,\ \forall f \in F
\end{aligned}
\tag{1}
$$

By the first two constraints, every integral feasible solution to LP (1) is a perfect matching on $W \cup F$ and by the third constraint, it has no blocking pairs. It is therefore a stable matching.

Taking α_w, β_f and γ_{wf} to be the dual variables for the first, second and third constraints of LP (1), respectively, we obtain the dual LP:

$$
\begin{aligned}
\min \quad & \sum_{w \in W} \alpha_w + \sum_{f \in F} \beta_f \\
\text{s.t.} \quad & \alpha_w + \beta_f + \sum_{f' \succ_w f} \gamma_{wf'} + \sum_{w \succ_f w'} \gamma_{w'f} \geq 1 \quad \forall w \in W,\ \forall f \in F, \\
& \gamma_{wf} \geq 0 \quad \forall w \in W,\ \forall f \in F
\end{aligned}
\tag{2}
$$

Lemma 62. *If x is a feasible solution to LP (1) then $\alpha = 0,\ \beta = 0,\ \gamma = x$ is an optimal solution to LP (2). Furthermore, if $x_{wf} > 0$ then*

$$\sum_{f \succ_w f'} x_{wf'} = \sum_{w' \succ_f w} x_{w'f}.$$

Proof. The feasibility of (α, β, γ) follows from the feasibility of x. The objective function value of this dual solution is 0, i.e., the same as that of the primal. Therefore, this solution is also optimal.

Since $\gamma = x$ is an optimal solution, if $x_{wf} > 0$ then $\gamma_{wf} > 0$. Now the desired equality follows by applying the complementary slackness condition to the third constraint of LP (1). □

Next, we will prove Theorem 65 and Corollary 66. Let x be a feasible solution to LP (1). Corresponding to x, we will define $2n$ unit intervals, I_w and I_f, one corresponding to each worker w and one corresponding to each firm f, as follows. For each worker w, we have $\sum_{f \in F} x_{wf} = 1$. Order the firms according to w's preference list; for simplicity, assume it is $f_1 \succ_w f_2 \succ_w \ldots \succ_w f_n$. Partition I_w into n ordered subintervals so that the ith interval has length x_{wf_i}; if this quantity is zero, the length of the interval is also zero. Next, for each firm f, we have $\sum_w x_{wf} = 1$. This time, order the workers according to f's preference list in reverse order; assume it is $w_1 \prec_f w_2 \prec_w \ldots \prec_w w_n$, and partition I_f into n ordered subintervals so that the ith interval has length $x_{w_i f}$.

Pick θ with uniform probability in the interval $[0, 1]$ and for each worker w, determine which subinterval of the interval I_w contains θ. The probability that any of these n subintervals is of zero length is zero, so we may assume that this event does not occur. Define perfect matching μ_θ as follows: if the subinterval of I_w containing θ corresponds to firm f, then define $\mu_\theta(w) = f$.

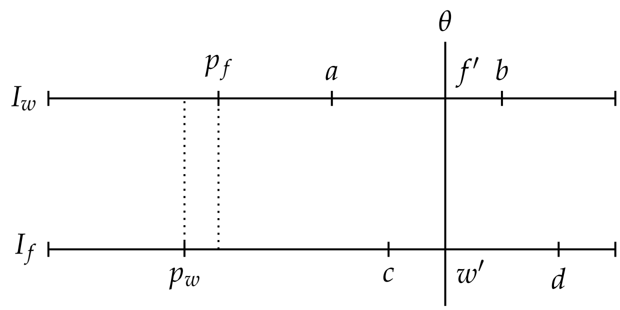


Figure 13

Lemma 63. *If $\mu_\theta(w) = f$ then the subinterval of I_f containing θ corresponds to worker w.*

Proof. Assume that the subinterval containing θ in I_w is $[a, b]$, where $a, b \in [0, 1]$; since it corresponding to f, $b - a = x_{wf}$. Since $x_{wf} > 0$, by the second part of Lemma 62, the subinterval corresponding to w in I_f is also $[a, b]$. Clearly, it contains θ. □

Lemma 64. *For each $\theta \in [0, 1]$, the perfect matching μ_θ is stable.*

Proof. Assume that $\mu_\theta(w) \neq f$. Let $\mu_\theta(w) = f'$ and $\mu_\theta(f) = w'$, where $f \succ_w f'$. We will show that $w' \succ_f w$, thereby proving that (w, f) is not a blocking pair. Extending this conclusion to all worker-firm pairs that are not matched by μ_θ, we get that μ_θ is a stable matching.

Since $\mu_\theta(w) = f'$, θ lies in the subinterval corresponding to f' in I_w. In Figure 13, this subinterval has been marked as $[a, b]$. θ also lies in subinterval corresponding to w' in I_f; this subinterval is marked as $[c, d]$ in Figure 13. Let p_f and p_w denote the larger endpoints of the subintervals containing f and w in I_w and I_f, respectively. By the third constraint of LP (1), the subinterval $[p_w, 1]$ of I_f is at least as large as the subinterval $[p_f, 1]$ of I_w. Since $f \succ_w f'$, θ lies in $[p_f, 1]$ and therefore, it also lies in $[p_w, 1]$. Therefore, $w' \succ_f w$ and the lemma follows. □

Theorem 65. *Every feasible solution to LP (1) is a convex combination of stable matchings.*

Proof. The feasible solution x of LP (1) has at most $O(n^2)$ positive variables x_{wf} and therefore there are $O(n^2)$ non-empty subintervals corresponding to these variables in the intervals I_w and I_f. Hence $[0, 1]$ can be partitioned into at most $O(n^2)$ non-empty subintervals, say $I_1, \ldots I_k$, so that none of these subintervals straddles the subintervals corresponding to the positive variables. For $1 \leq i \leq k$, each $\theta \in I_i$ corresponds to the same stable matching $\mu_i = \mu_\theta$. Let the length of I_i be α_i; clearly $\sum_{i=1}^{k} \alpha_i = 1$. Then we have

$$x = \sum_{i=1}^{k} \alpha_i \mu_i.$$

The theorem follows. □

Corollary 66. *The polyhedron defined by the constraints of LP (1) has integral optimal vertices; these are stable matchings.*

5.2 LPs for Settings II and III

For Settings II by the Rural Hospital Theorem 23, the set of workers and firms matched in all stable matchings is the same. Let these sets be W' and F', respectively. Clearly, it suffices to work with the primal LP (1) restricted to these sets only.

For Setting III, we will use the reduction from Setting III to Setting II given in Section 2.3.

6 Exercises

1. In Setting I, say that a perfect matching μ is *Pareto optimal* if there is no perfect matching $\mu' \neq \mu$ under which each agent is weakly better off, i.e., for each agent $a \in W \cup F$, $\mu(a) \succeq_a \mu'(a)$. Prove that every stable matching is Pareto optimal. Is the converse true? Give a proof or a counter-example.

2. Define a *preferred couple* in Setting I to be a worker-firm pair such that each is the most preferred in the other's preference list.

- Prove that if an instance has a preferred couple (w, f), then in any stable matching, w and f must be matched to each other.
- Prove that if in an instance with n workers and n firms there are n disjoint preferred couples, then there is a unique stable matching.
- Prove that the converse of the previous problem does not hold, i.e., there is an instance which does not have n disjoint preferred couples, yet it has a unique stable marriage.

3. Construct an instance for Setting I that has exponentially many stable matchings.

4. Design a polynomial time algorithm which given an instance for Setting I, finds an unstable perfect matching, i.e., one having a blocking pair.

5. In Setting II, let us add a second termination condition to the algorithm, namely terminate when every firm gets at least one proposal. If this condition holds, match every firm to its best proposal, leaving the rest of the workers unmatched. Give a counter-example to show that this matching is not stable.

6. These exercises are on the issue of incentive compatibility.

1. Find an instance $(W, F, \succ)$ in which some worker is matched to the last firm with respect to its preference list in the worker-optimal stable matching. How many such workers can there be in a instance?
2. For Setting I, find an instance in which a firm can improve its match in the worker-optimal stable matching by misreporting its preference list.

3. For Setting I, find an instance in which two workers can collude, i.e. lie together, in such a way that one of them does better and the other does no worse, in the worker-optimal stable matching.

7. Use the fact that the operations of min and max on two element sets satisfy the distributive property to prove that the operations of meet and join for the lattice of stable matching also satisfy this property.

8. ([Bla84]) Prove that lattices arising from instances of stable matching form a *complete set of FDLs*, i.e., every FDL is isomorphic to some stable matching lattice.

9. These exercises are on rotations.

1. Give an example of a stable matching μ such that $\text{next}(w) \neq \perp$ and yet w is not in a rotation with respect to μ.
2. Let ρ be a rotation with respect to stable matching μ. Obtain ρ' by permuting the sequence of pairs in ρ. Show that if the permutation is a cyclic shift then $\rho'(\mu) = \rho(\mu)$ and otherwise the matching $\rho'(\mu)$ is not stable.
3. Let ρ be a rotation with respect to stable matching μ and let $\rho(\mu) = \mu'$. Prove that μ' is a direct successor of μ, i.e., $\mu \rhd \mu'$.
4. Show that any path in G_μ from $\mu_\top$ to $\mu_\perp$ involves applying each rotation in $\mathcal{R}_\mu$ exactly once. Use this fact to give a polynomial time algorithm for finding all rotations in $\mathcal{R}_\mu$.
5. Give a polynomial time algorithm for: Given a rotation $\rho \in \mathcal{R}_\mu$, find the join-irreducible it corresponds to.

10. These exercises are on Birkhoff's Theorem.

1. Give a polynomial time algorithm for computing the succinct description of lattice $\mathcal{L}_\mu$ promised in Corollary 53, i.e., π_μ.
2. Let S be a lower set of π_μ and let S_ρ be the set of rotations corresponding to the join-irreducibles in S. Let τ be any topological sort of the rotations in S_ρ that is consistent with the partial order π_μ. Show that if starting with the matching $\mu_\top$, the rotations in S_ρ are applied in the order given by τ, then the matching obtained in $\mathcal{L}_\mu$ will be $f_\mu^{-1}(S)$. Use this fact to show that $f_\mu^{-1}(S)$ can be computed in polynomial time for any lower set S of π_μ.

11. For any positive integer n, let S_n denote the set of divisors of n. Define partial order $\pi_n = (S_n, \succeq)$ as follows: for $a, b \in S_n$, $a \succeq b$ if $b|a$. Prove that π_n is a FDL with the meet and join of two elements $a, b \in S_n$ being the gcd and lcm of a and b, respectively. Figure 14 shows the lattice for $n = 60$. Give a characterization of the join-irreducibles of π_{60} and find the projection of π_{60} onto the join-irreducibles of this lattice. Do the set of lattices $\{\pi_n \mid n \in \mathbb{Z}_+\}$ form a complete set of FDLs, as defined in Exercise 8? Prove or disprove.

12. ([ILG87]) For a stable matching μ in Setting I, define its *value* as follows: Assume $\mu(w) = f$, and that f is the j^{th} firm on w's preference list and w is the k^{th} worker on f's list; if so, define the value of the match (w, f) to be $k + j$. Define the value of μ to be the sum of values of all matches

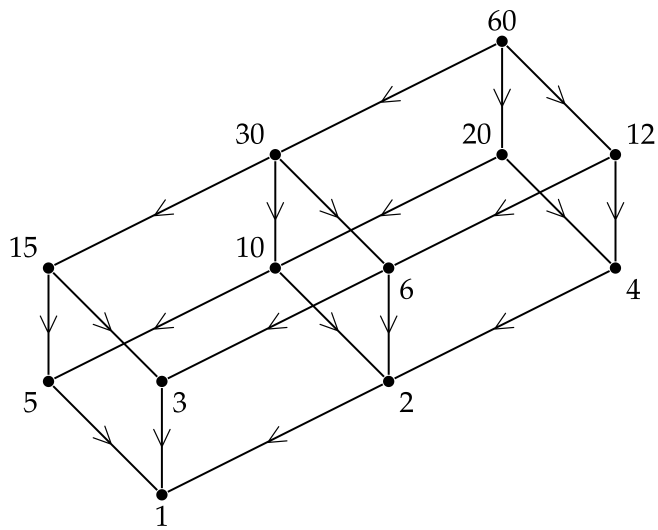


Figure 14: The lattice of divisors of 60.

in μ. Define an *eqitable stable matching* to be one that minimizes the value. Give a polynomial time algorithm for finding such a matching.

Hint: Use the algorithms developed in Exercise 10. Also note that the problem of finding a minimum weight lower set of π_μ is solvable in polynomial time, assuming that integral weights (positive, negative or zero) are assigned to the elements of $\mathcal{J}_\mu$. The weight of a lower set S is defined to be the sum of weights of elements in S.

13. Prove Lemma 59 and use it to prove Lemma 60.

Hint: For the first part of Lemma 59, use a blocking pair argument, and for the second part, first prove

$$|W_1| + |W_2| = \sum_{f \in F_1} n_1(f) + \sum_{f \in F_2} n_2(f),$$

where $n_1(f) = |\mu_1(f) - \mu_2(f)|$ and $n_2(f) = |\mu_2(f) - \mu_1(f)|$.

14. LP (1) for the stable matching problem was derived from a sufficient condition, given in Section 5.1, which ensures that a worker-firm pair (w, f) does not form a blocking pair with respect to a matching μ. Another sufficient condition is that if f is matched to firm w' such that $w \succ_f w'$ then w should be matched to a worker f' such that $f' \succ_w f$. The fractional version of this condition is:

$$\forall w \in W, \; \forall f \in F : \quad \sum_{w \succ_f w'} x_{w'f} - \sum_{f' \succ_w f} x_{wf'} \leq 0$$

Show that this condition holds for any feasible solution x to LP (1).

15. ([TS98]) Suppose that an instance of stable matching in Setting I has an odd number, k, of stable matchings. For each worker w, order its k matches, with multiplicity, per its preference list and do the same for each firm f. Match w to the median element in its list. Let us call this the *median matching*. This exercise eventually helps show that not only is this matching perfect, but it is also stable. Moreover, it matches each firm to the median element in its list as well.

First, let $\mu_1, \ldots \mu_l$ be any l stable matchings, not necessarily distinct, for an instance of stable matching in Setting I. For each worker-firm pair (w, f) let $n(w, f)$ be the number of these matchings in which w is matched to f and let $x_{wf} = (1/l) \cdot n(w, f)$.

Show that x is a feasible solution to LP (1), i.e., it is a fractional stable matching. For any k such that $1 \leq k \leq l$, let $\theta = (k/l) - \epsilon$, where $\epsilon > 0$ is smaller than $1/l$. Consider the stable matching μ_θ as defined by the procedure given in Section 5.1 for writing a fractional stable matching as a convex combination of stable matchings. Show that matching μ_θ matches each worker w to the kth firm in the ordered list of the l firms, not necessarily distinct, which w is matched to under $\mu_1, \ldots, \mu_l$. Furthermore, show that μ_θ matches each firm f to the $(l - k + 1)$th worker in the ordered list of the l workers that f is matched to.

Using the above-stated fact, prove the assertions made above about the median matching.

7 Notes

The seminal paper of Gale and Shapley [GS62] introduced the stable matching problem and gave the Deferred Acceptance Algorithm. For basic books on this problem and related topics see Gusfield and Irving [GI89], Roth and Sotomayor [RS92], Knuth [Knu97], and Manlove [Man13].

Dubins and Freedman [DF81] proved that the worker-proposing DA Algorithm is DSIC for workers. This ground-breaking result was instrumental in opening up the DA Algorithm to highly consequential applications, such as school choice; for a discussion of the latter, see Chapter **??**. Theorem 35, showing that there is no DSIC mechanism for firms in Setting III, is due to Roth[Rot85].

John Conway proved that the set of stable matchings of an instance forms a finite distributive lattice, see [Knu97]. The notion of rotation is due to Irving and Leather [IL86] and Theorem 56 is due to Birkhoff [Bir+37]. The LP formulation for stable matching was given by Vande Vate [Vat89]; the proof given in Section 5.1 is due to Teo and Sethuraman [TS98].

References

[Bir+37] Garrett Birkhoff et al. "Rings of sets". In: *Duke Mathematical Journal* 3.3 (1937), pp. 443–454.

[Bla84] Charles Blair. "Every finite distributive lattice is a set of stable matchings". In: *Journal of Combinatorial Theory, Series A* 37.3 (1984), pp. 353–356.

[DF81] L. E. Dubins and D. A. Freedman. "Machiavelli and the Gale-Shapley Algorithm". In: *The American Mathematical Monthly* 88.7 (1981), pp. 485–494. ISSN: 00029890, 19300972. URL: http://www.jstor.org/stable/2321753.

[GI89] Dan Gusfield and Robert W Irving. *The stable marriage problem: structure and algorithms*. MIT press, 1989.

[GS62] David Gale and Lloyd S Shapley. "College admissions and the stability of marriage". In: *The American Mathematical Monthly* 69.1 (1962), pp. 9–15.

[IL86] Robert W Irving and Paul Leather. "The complexity of counting stable marriages". In: *SIAM Journal on Computing* 15.3 (1986), pp. 655–667.

[ILG87] Robert W Irving, Paul Leather, and Dan Gusfield. "An efficient algorithm for the "optimal" stable marriage". In: *Journal of the ACM (JACM)* 34.3 (1987), pp. 532–543.

[Knu97] Donald Ervin Knuth. *Stable marriage and its relation to other combinatorial problems: An introduction to the mathematical analysis of algorithms*. American Mathematical Soc., 1997.

[Man13] David Manlove. *Algorithmics of Matching Under Preferences*. World Scientific, 2013.

[Rot85] A. E. Roth. "The college admissions problem is not equivalent to the marriage problem". In: *Journal of Economic Theory* 36.2 (1985), pp. 277–288. ISSN: 0022-0531. DOI: https://doi.org/10.1016/0022-0531(85)90106-1. URL: http://www.sciencedirect.com/science/article/pii/0022053185901061.

[RS92] Alvin E. Roth and Marilda Sotomayor. *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis*. Econometric Society Monographs, 1992.

[TS98] Chung-Piaw Teo and Jay Sethuraman. "The geometry of fractional stable matchings and its applications". In: *Mathematics of Operations Research* 23.4 (1998), pp. 874–891.

[Vat89] John H Vande Vate. "Linear programming brings marital bliss". In: *Operations Research Letters* 8.3 (1989), pp. 147–153.