CS 6002: Selected Areas of Mechanism Design **Jan-Apr 2025**

Lecture 5: Structures of Stable Matchings

Lecturer: Swaprava Nath *Scribe(s): Shresth Verma, Ananya Kulashreshtha*

Disclaimer: *These notes aggregate content from several texts and have not been subjected to the usual scrutiny deserved by formal publications. If you find errors, please bring them to the notice of the Instructor.*

5.1 Introduction

In this lecture, we explore the concept of men-optimal matchings within the framework of stable matching theory. We will delve into the properties of men-optimal and women-optimal matchings, the Gale-Shapley Deferred Acceptance (DA) algorithm, and the lattice structure of stable matchings.

5.2 Preliminaries and Definitions

Definition 5.2.1 (Achievable)**.** *A man m and a woman w are achievable for each other if there exists some stable matching where they are paired together.*

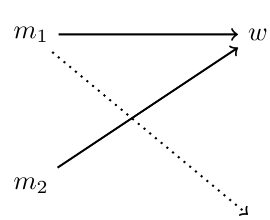
Favorite Achievable Partner: Because of strict preferences, among all the achievable men/women of a woman/man (which can appear in different stable matches), there exists **exactly one** favorite achievable man/woman.

Definition 5.2.2 (Men-Optimal Function)**.** *Consider a function $f^M : M \to W$ that maps every man to his most preferred (favorite) achievable woman. This function is called the men-optimal function.*

Definition 5.2.3 (Women-optimal Function)**.** *Similarly, define $f^W : W \to M$ as the function that maps every woman to her most preferred (favorite) achievable man. This is known as the women-optimal function.*

Theorem 5.2.4. *The men-optimal function returns a matching. In other words, the function $f^M : M \to W$ is a bijection (Analogus for women-optimal).*

Proof. Assume, for contradiction, that f^M is not a matching. Then there exist two distinct men m_1 and m_2 such that $f^M(m_1) = f^M(m_2) = w$ for some woman w.



Since w is the favorite achievable woman for both m_1 and m_2, consider their preferences. Suppose w prefers m_1 over m_2, i.e., $P_w(m_1) > P_w(m_2)$.

By the definition of achievability, there exists a stable matching μ where m_2 is matched with w. In this matching μ, m_1 must be matched to some woman w' whom he prefers less than w (since w is his favorite achievable woman). However, this creates a blocking pair (m_1, w) because:

1. m_1 prefers w over w', and
2. w prefers m_1 over m_2.

This contradicts the stability of μ. Therefore, f^M must be a matching. □

5.3 The Gale-Shapley Deferred Acceptance Algorithm

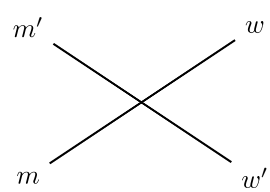
The Gale-Shapley Deferred Acceptance (DA) algorithm is a foundational method for finding stable matchings. We had discussed the man-proposing version of this algorithm and now will prove its optimality.

Theorem 5.3.1. *For every preference profile P, the matching computed by the men-proposing DA algorithm is the* ***men-optimal*** *stable matching.*

Proof. **Goal:** To show that in the men-proposing DA algorithm, no man is ever rejected by his favorite achievable woman.

Proof by Contradiction:

1. **Assumption:** Suppose there exists a man m who is the first to be rejected by his favorite achievable woman w during the DA process.
2. **Implications for w:** If w rejects m, it must be because w has received a proposal from another man m', whom she prefers over m (i.e., $P_w(m') > P_w(m)$).



3. **Behavior of m':** When m' proposes to w, any rejections m' received in the past must have been from women who are unachievable for him. (This is because m is assumed to be the first man rejected by an achievable woman.)
4. **Existence of a Stable Matching μ:** By the definition of achievability, there exists a stable matching μ where m and w are matched. Since m and w are achievable for each other, their pairing must be part of μ.
5. **Implication for m':** Under μ, m' must be matched with a woman w' whom he prefers less than w, because all women he prefers more than w are unachievable for m'.
6. **Contradiction in Stability:** The pair (m', w) forms a blocking pair in μ because:
   (a) m' prefers w over his partner in μ, and
   (b) w prefers m' over her partner in μ.

This violates the stability of μ, contradicting the assumption that μ is a stable matching.

7. **Conclusion:** Therefore, our initial assumption is false. No man is rejected by his favorite achievable woman during the DA process, and the algorithm produces the **men-optimal** stable matching.

□

5.4 Knuth's Theorem on Stable Matchings

Theorem 5.4.1. *For any distinct stable matchings μ and μ', if all men prefer μ at least as good as μ', then all women prefer μ' at least as good as μ.*

Proof. (by Contradiction)

1. **Assumption:** Suppose there exists a woman w who prefers μ over μ'.
2. **Partners in the Matchings:** Let m be w's partner in μ, and let m' be w's partner in μ'. Thus, the matchings are:

3. **Preference of m:** Since m prefers μ over μ', we have $w\,P_m^{\mu'}\,w'$. (This means m prefers w over his partner w' in μ'.)
4. **Preference of w:** Since w prefers μ over μ', we have $m\,P_w^{\mu'}\,m'$. (This means w prefers m over her partner m' in μ'.)
5. **Blocking Pair:** The pair (m, w) forms a blocking pair in μ' because:
   (a) m prefers w over his partner in μ', and
   (b) w prefers m over her partner in μ'.

   This violates the stability of μ', leading to a contradiction.
6. **Conclusion:** Our assumption is false. If all men prefer μ over μ', then all women must prefer μ' over μ.

□

Discussion: The above theorem establishes a relationship between men's and women's preferences under stable matchings when a clear hierarchy of preferences exists. However, it does not directly address what happens when the matchings μ and μ' are incomparable.

Resolving Incomparable Matchings: We can address such cases by constructing new matchings (which are also stable) where there is consensus among men and women. The process is as follows:

1. Start with an incomparable pair of matchings μ and μ'.
2. Construct a new matching where:
   - All men prefer more (and women prefer less), or
   - All men prefer less (and women prefer more).

5.5 Mapping Between Stable Matchings: The Function $\max_{P,Q}$

Let P and Q be any pair of stable matchings.

5.5.1 Definition of $\max_{P,Q}$

Define the mapping $\max_{P,Q}$ such that:

1. Each man maps to his **more preferred woman** between P and Q.
2. Each woman maps to her **less preferred man** between P and Q.

Using the same function from both ends:

$$\max_{P,Q}(m) \quad \text{and} \quad \max_{P,Q}(w)$$

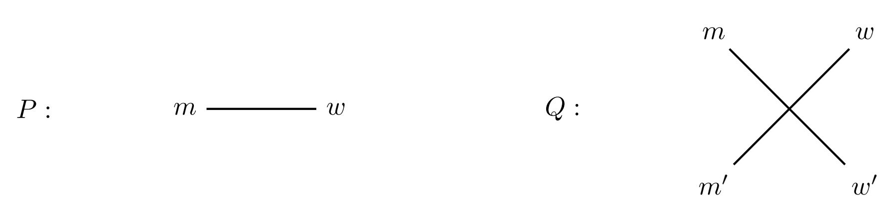
Lemma 5.5.1. *$max_{P,Q}$ yields a matching.*

Proof. It suffices to show that for any pair of m and w:

$$\max_{P,Q}(m) = w \iff \max_{P,Q}(w) = m.$$

(⇒) Direction:

- Suppose not, i.e., $\max_{P,Q}(m) = w$ but $\max_{P,Q}(w) = m' \neq m$.
- Since $\max_{P,Q}(m) = w$, there must be one matching between P and Q where m is matched to w. Say it is P:



- Since $\max_{P,Q}(w) = m' \neq m$, then m' has to be below m in w's preference (by definition of $\max_{P,Q}$).
- Thus, Q is not stable as (m, w) is a blocking pair.

(⇐) Direction: Using the first part, we can claim that two distinct men cannot point to the same woman:

- Since that woman will point back to exactly one man, $\max_{P,Q}$ is a well-defined function.
- Also, each man points to a woman. Hence it has to be distinct.
- Since the cardinality $|M| = |W|$, the mapping $\max_{P,Q}$ must be a matching.

Alternative Way of Thinking:

- $\max_{P,Q}$ is a mapping from $M \to W$ (from the men's side).
- The first part shows that it is 1-1 (injective).
- Since a mapping between two finite sets of equal cardinality must be bijective, the function must be a bijection (hence a matching).

$\square$

5.5.2 Use: Incomparable Stable Matchings

- Incomparable stable matchings imply no consensus.
- However, all men/women can immediately construct a different stable matching that they can agree on.

$$\begin{tikzcd}
 & \max_{P,Q} & \\
 & (1,1,2,2) \arrow[dl] \arrow[dr] & \\
(1,2,3,2) \arrow[dr] & & (2,1,2,3) \arrow[dl] \\
 & (2,2,3,3) & \\
 & \min_{P,Q} &
\end{tikzcd}$$

Figure 5.1: Construction of $\max_{P,Q}$ and $\min_{P,Q}$ from incomparable stable matchings

5.5.3 Extending $\max_{P,Q}$ and $\min_{P,Q}$

Theorem 5.5.2 (Lattice Theorem)**.** *The mappings $max_{P,Q}$ and $min_{P,Q}$ induce stable matchings.*

Lemma 5.5.3. *The mapping $max_{P,Q}$ yields a stable matching.*

Proof. Suppose, for contradiction, that (m, w) blocks $\max_{P,Q}$.

Since m prefers w over $\max_{P,Q}(m) = w_1$ (say), w_1 is the more preferred woman for m in P and Q. Hence, w will be even above w_1 in P_m and Q_m:

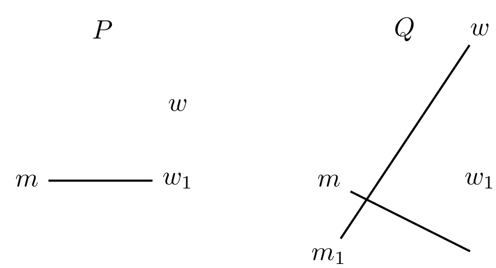
$$w \, P_m \max_{P,Q}(m).$$

Now consider w. Suppose w prefers P over Q. Hence, $\max_{P,Q}(w)$ is her Q-matching (according to the definition of $\max_{P,Q}(w)$). Let the man matched to w in Q be m_1. We claim:

m_1 is below m in w's preferences.

Otherwise, (m, w) cannot be a blocking pair of $\max_{P,Q}$. If m_1 which is the worse match between P and Q's matching of w is above m in w's preference, then (m, w) can't make a blocking pair of $\max_{P,Q}$.

$$m \, P_w \max_{P,Q}(w).$$



Regardless of which matching gives the worse match for w, that matching is blocked by (m, w).

Since both P and Q are stable matchings, this is a contradiction. Hence $\max_{P,Q}$ yields a stable matching.

□

5.5.4 Extension to $\min_{P,Q}$

Similarly, the $\min_{P,Q}$ mapping can be defined as a mirror opposite from the woman's side. Following similar arguments, we can show that $\min_{P,Q}$ is also a stable matching.

Corollary 5.5.4. *The mapping $\min_{P,Q}$ yields a stable matching.*

5.6 Conclusion

The mappings $\max_{P,Q}$ and $\min_{P,Q}$ provide powerful tools for constructing new stable matchings from incomparable stable matchings. The lattice structure of stable matchings guarantees that such constructions are both consistent and stable. These mappings highlight the structural richness of stable matching theory and its practical applications.