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

Lecture 4: Two Sided Matching

Lecturer: Swaprava Nath *Scribe(s): Vijay*

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 to the notice of the Instructor.*

4.1 Two Sided Matching

A market is said to be **two-sided** if there are two sets of agents, and if an agent from one side of the market can be matched only with an agent from the other side, and each agent has a preference over the agents to be matched with from the other group. **Two sided matching** is a pairing-up of agents from one side with agents of the other side.

In this lecture, we restrict ourselves to balanced markets, i.e., both the sides of the market have the same number of agents.

The motivation to study two-sided matchings is evident: the main goal of several economic constructs like markets, and social processes is to match one type of agent with another. This includes matching students to colleges, doctors to hospitals and marriageable men and women.

4.2 Gale and Shapley's Deferred Acceptance Algorithm

In 1962, a class of "two-sided matching models" were introduced by Gale and Shapley [GSDA] to study college admissions and marriages.

In these notes, we will now assume that there is a set of men M, against an equally sized set W of women, and each man wishes to find a match from the set of women and vice versa.

4.2.1 The Algorithm

The key idea of the algorithm (hereafter, referred to as DA) [1] can be summarized in one line: *"Let every man propose to that woman who he prefers the most, out of the women who have not yet rejected him"*.

[1] mpDA referes to the algorithm when men constitute the proposing set of agents, and wpDA when women constitute the proposing set of agents. DA in general refers to mpDA here.

Algorithm 1: Gale and Shapley's Deferred Acceptance Algorithm

Input: Two sets: M (men) and W (women), with preferences for each agent over the other set.
Output: A stable matching between M and W.
Initialize each $m \in M$ and $w \in W$ as unmatched.
Initialize preference lists for all agents.
while *there exists an unmatched* $m \in M$ **do**
    m proposes to the most preferred w who has not rejected him yet.
    if w *is unmatched* **then**
        (m, w) becomes a tentative match.
    else
        Let m' be w's current match.
        if w *prefers* m *over* m' **then**
            (m, w) becomes a tentative match.
            m' becomes free.
        else
            m remains free.
Finalize all tentative matches.

Having stated the algorithm, the following illustration demonstrates a simple execution of the algorithm.

4.2.2 Illustration of DA

Consider the following example (from the animation presented in class) of a two sided market with 4 individuals on each side.

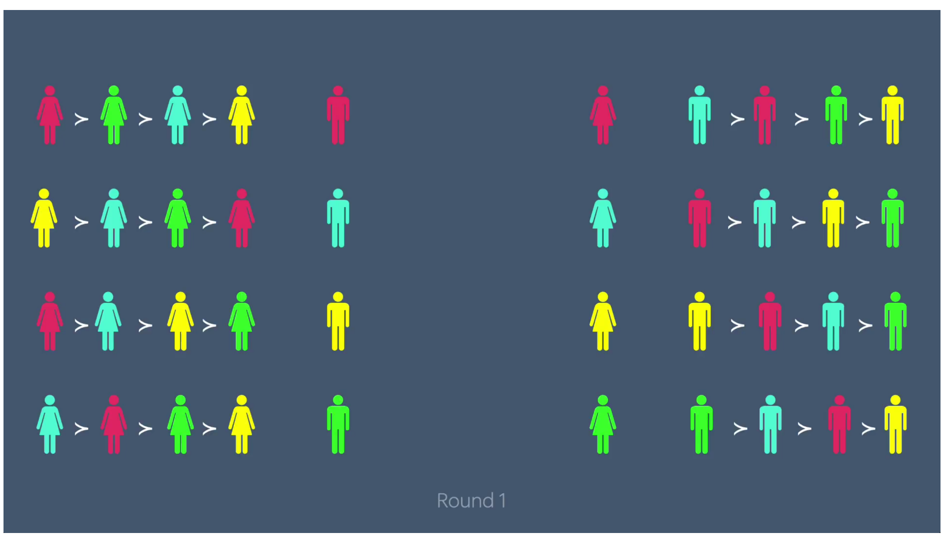

Figure 4.1: Initial setup

Initial stage: The initial setup is shown in Fig 1. The preference order of every agent is listed along their sides.

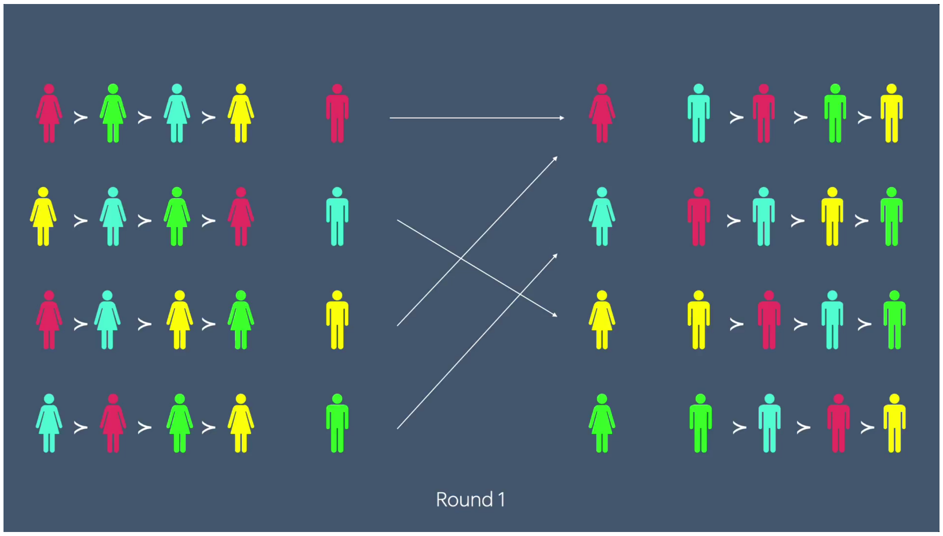

Figure 4.2: Round 1 proposals

Round 1 proposals: Each male proposes to their most preferred female. w_{red} has received two proposals and w_{green} received none.

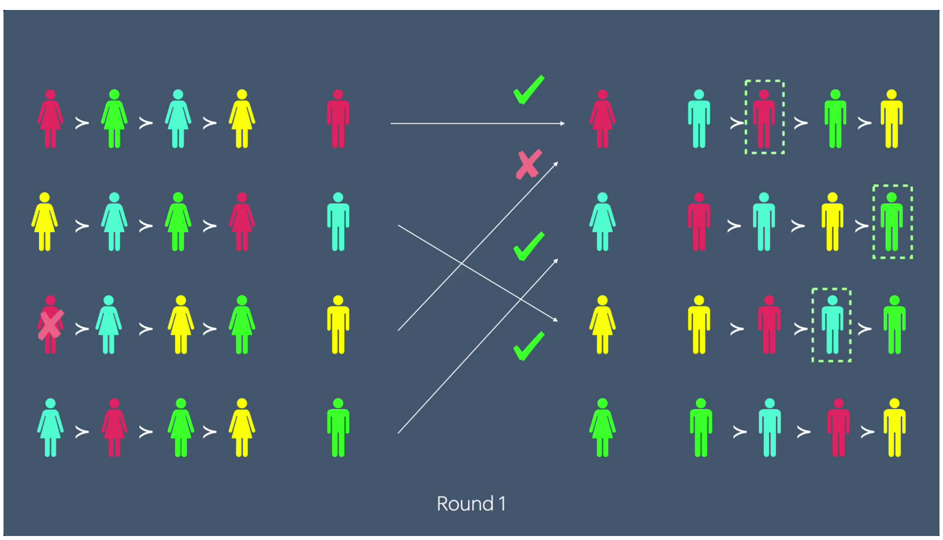

Figure 4.3: Round 1 acceptances and rejection

Round 1 proceedings: Women with only one proposal accept the offer. w_{red} accepts m_{red}'s proposal because she prefers him more over m_{yellow}.

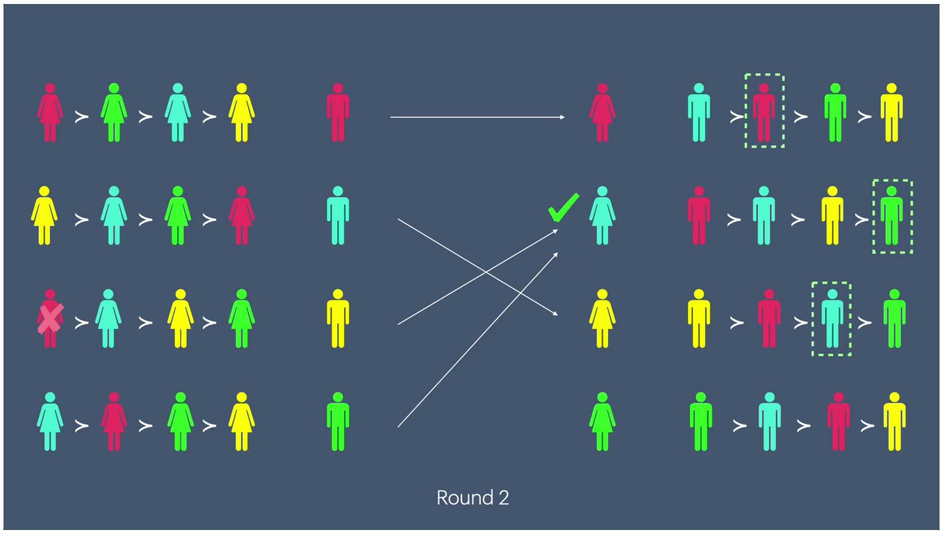

Figure 4.4: Round 2

Round 2: Only m_{yellow} is unmatched. So he proposes to his next top preference who has not rejected him. This is w_{blue}. w_{blue} accepts this offer as she prefers him over her tentative match with m_{green}. m_{green} now is unmatched.

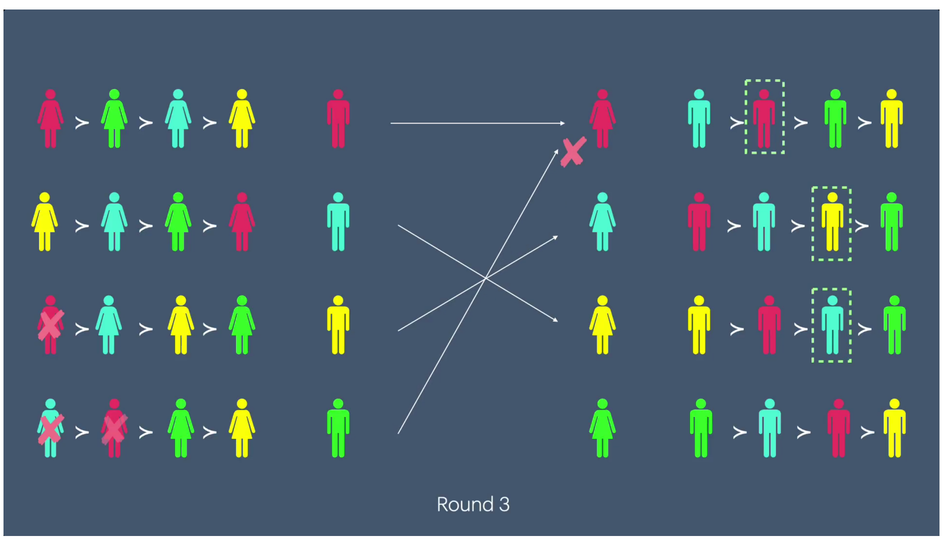

Figure 4.5: Round 3

Round 3: Only m_{green} is unmatched. So he proposes to his next top preference who has not rejected him. This is w_{red}. w_{red} rejects this offer as she prefers him less than her tentative match, m_{red}. m_{green} remains unmatched.

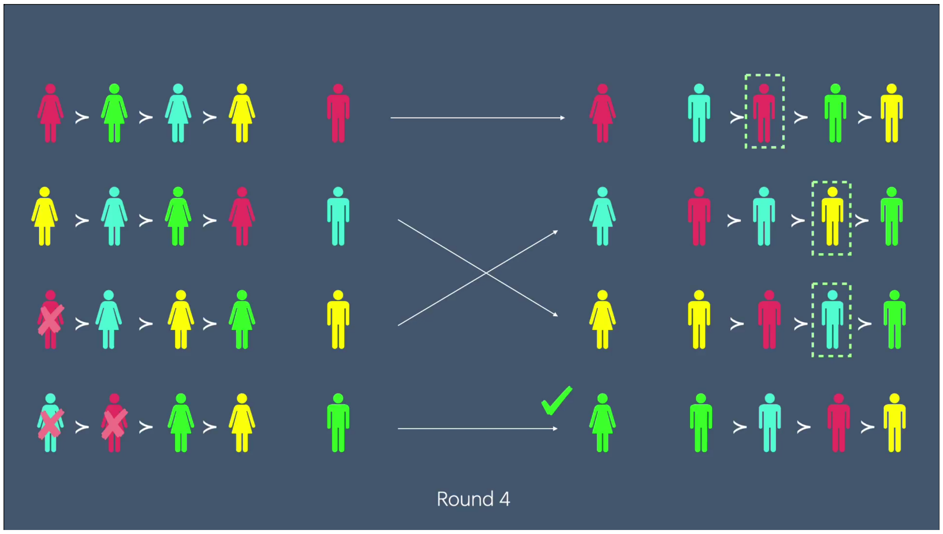

Figure 4.6: Round 4

Round 4: Only m_{green} is unmatched. So he proposes to his next top preference who has not rejected him. This is w_{green}. w_{green} accepts this offer as she has not yet found a match.

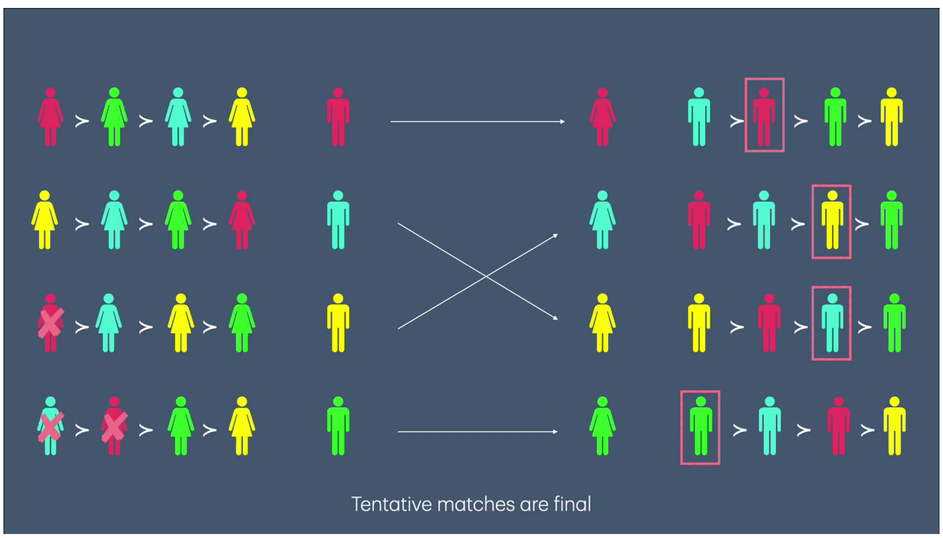

Figure 4.7: Termination of DA

Termination: Everyone is now matched and has found a match. All tentative matches become final and DA terminates.

4.3 Properties of DA

Claim 4.1 *DA terminates with convergence in polynomial-time*

Proof: Any man will propose to each woman at most once, since DA prevents proposals to the same woman. Thus any man may make at most n requests, where n is the number of women, and there are n such men proposing in all. No man proposes a woman who rejected him in the previous round. Hence no proposals are repeated. Therefore the algorithm must terminate in $\mathcal{O}(n^2)$ time. ■

Lemma 4.2 *There is always at least one woman at the termination of DA, who did not receive more than 1 proposal.*

Proof: Suppose that every woman gets two or more proposals. Then in the very last round of DA, there might be a woman who:

- who receives multiple proposals. If so, such a woman will reject one of the proposals, leaving a man unmatched.
- who receives only one proposal, but may have accepted an earlier proposal. This means she already has a tentative match. No matter which man she chooses finally, she leaves one of the men unmatched.

By the very definition of DA, loop termination at the final round implies that there are no unmatched men. Thus, we arrive at a contradiction. Our assumption is wrong. ■

Corollary 4.3 *From the above lemma, the last round will always have a woman who has received a single proposal for the first time.*

Claim 4.4 *At most $1 + n(n-1)$ total proposals are made in DA, and this is the least upper bound on total proposals.*

Proof: From the previous corollary, the last round has a woman who received one proposal for the first time. Other than her, the $n-1$ women may have possibly reached n proposals, one from each man. Thus, the total number of proposals made in the DA is limited to $1 + n(n-1)$.

Now consider the following example where this is actually the number of proposals. This will conclude that this is in fact the least upper bound on the number of proposals.

Example:
Consider a marriage market with 3 men (m_{blue}, m_{black} and m_{yellow}) and 3 women (w_{green}, w_{grey} and w_{blue}) with the preferences depicted by the image below.

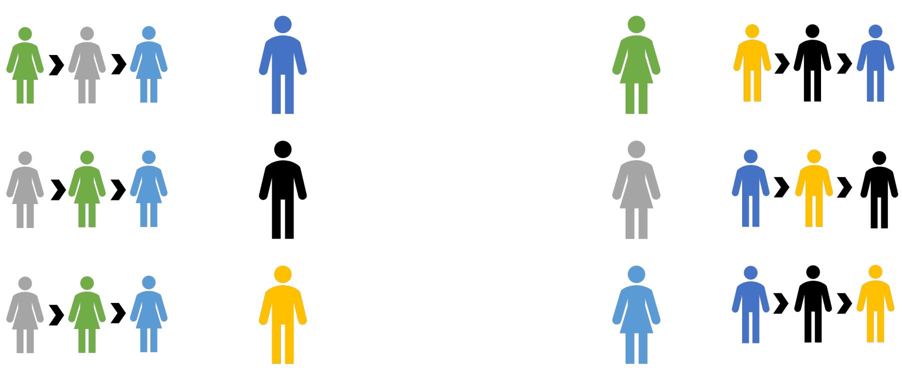

Figure 4.8: A 3-sized marriage market with 7 proposals

We now proceed to show that there are indeed 1+3(3-1) proposals in a correct execution of DA in the following scenario. This would prove our claim.

1. The algorithm begins with no one being matched. m_{blue} proposes to w_{green}, which she accepts. There is a tentative match.
2. Unmatched m_{black} proposes to w_{grey}, who accepts the proposal.
3. Unmatched m_{yellow} proposes to w_{grey}. She prefers m_{yellow} over m_{black} and accepts the proposal leaving m_{black} unmatched.
4. m_{black} proposes to w_{green}, which she accepts, thereby leaving m_{blue} unmatched.
5. m_{blue} proposes to w_{grey}. Since he is her highest preference, she will leave m_{yellow} for him.
6. Unmatched m_{yellow} approaches his next non-rejecting preference, w_{green}. She accepts the proposal since he is her favourite preference and leaves m_{black} unmatched.
7. m_{black} has been rejected by the others, and so he proposes to w_{blue} who accepts this first offer she received. Everyone is matched and DA terminates.

This enumerates 7 rounds. Thus, $n(n-1)+1$ constitutes the tightest upper bound on the number of rounds of execution by DA. ■

Claim 4.5 *DA returns a perfect matching.*

Proof: A woman will always reject her previous proposal to match a new man, therefore all women are matched with only one man at most. A man will only approach a woman and propose if he is unmatched, therefore, every man is at most matched with one woman at the end of DA. DA terminates when everyone has at least one match.

Therefore, DA returns a perfect matching since everyone finds exactly one match. ■

4.4 Stability and Coalitions in DA

4.4.1 Notation and Definitions

Definition 4.6 *A **Matching** (denoted $\mu : M \to W$) is a bijective map from the set of men M to the set of women W.*

Under a given μ we say that $\mu(m)$ denotes the woman paired with m under this mapping and with some notational shorthand we denote by $\mu(w)$ (which really should be $\mu^{-1}(w)$) the man paired with the woman w. Let also P denote a preference profile of strict preferences of each agent over the other agents and let the set of all such profiles be $\mathcal{P}$.

Definition 4.7 *A matching μ is said to be **pairwise unstable** if there exists a preference profile P for which there exists a pair (m, w) with $m \in M, w \in W$, such that:*

$$w \; P_m \; \mu(m) \text{ and } m \; P_w \; \mu(w) \tag{4.1}$$

*and we say that (m, w) constitutes a **blocking pair** of μ given P.*

Definition 4.8 *If a matching μ^* has no blocking pair for $P \in \mathcal{P}$, then μ^* is **pairwise stable** at P*

Note: For brevity, saying **"μ is pairwise stable"** without any reference to a preference profile, would indicate that it is pairwise stable at all $P \in \mathcal{P}$.

Definition 4.9 *For a given matching μ, a coalition subset $S \subseteq M \cup W$ is said to be a* ***blocking group*** *or* ***blocking coalition*** *given at some preference profile $P \in \mathcal{P}$, if another matching μ' can be contrived such that:*

$$\begin{aligned} &(i)\ \forall w \in S \cap W,\ \ \mu'^{-1}(w) \in S \cap M \\ &\quad\ \forall m \in S \cap M,\ \ \mu'(m) \in S \cap W \\ &(ii)\ \mu'(m)\ P_m\ \mu(m),\ \forall m \in S \cap M \\ &\quad\ \mu'^{-1}(w)\ P_w\ \mu^{-1}(w),\ \forall w \in S \cap W \end{aligned}$$

Definition 4.10 *A matching μ is said to be in the* ***core*** *of profile $P \in \mathcal{P}$, if there is no blocking coalition at μ given P.*

4.4.2 Stability Theorems in DA

Claim 4.11 *For any profile $P \in \mathcal{P}$, DA returns a pairwise stable matching.*

Proof: Suppose DA is not pairwise stable at P. By definition of pairwise stability, this means that there exists a blocking pair (m, w) such that $w\ P_m\ \mu(m)$ and $m\ P_w\ \mu(w)$.

In DA, m could have only been matched to someone he prefers less than w if he was rejected by w, since men will choose their highest non-rejecting choice to propose to. But if w did in fact reject m, she must have done it for a more preferred man, by definition of DA. This directly contradicts our assumption. ■

The proof of the above claim directly leads to the following corollary.

Corollary 4.12 *In DA, after each round/proposal, the man that any woman is matched with (that is, if she is matched at all), is either as preferred or more preferred than any men she was matched with in the previous rounds.*

Now we come to a very important result.

Theorem 4.13 *A matching that is pairwise stable at a preference profile $P \in \mathcal{P}$ if and only if the matching is in the core of that profile.*

Proof:
(Core $\implies$ Pairwise stability)
It is easy to see that given a matching μ in a preference profile P that is in the core of P, there are no blocking coalitions S of any size. This includes the blocking coalitions with one man and one woman, which means there are no blocking pairs as well.

(Pairwise stability $\implies$ Core)
We prove the contrapositive (Not in Core $\implies$ No Pairwise stability). Suppose that μ is not in the core of P. This means that, there exists a blocking coalition S, and a new preferred matching μ' by the members

of S. Let m be one of the men in M. Under the new matching μ'.

The "for all" condition (ii) in the definition of a blocking coalition implies that both m and his new match $w_1 := \mu'(m)$ are better off under this new matching.
Thus, $w_1 \; P_m \; \mu(m)$ and $m \; P_w \; \mu^{-1}(w_1)$.

Irrespective of the rest of the members in the coalition, m and w_1 could have just deviated from μ by choosing each other. Therefore, they form a blocking pair. Thus whenever a blocking coalition exists, a blocking pair also exists. The contrapositive is proven. ■

This gives us a unified concept of "stability", since group and pairwise stability are provably equivalent notions. We say that DA gives us a "stable matching".

4.5 Lattices of Stable Matchings

Notice that it is possible to have different stable matchings. For a given marriage market M, W, and a preference profile P, *wpDA* and *mpDA* will produce different matchings. Both of these matchings would be stable.

Consider the two sided market given below with the preferences of each individual as depicted. Before we consider the possibility of multiple matchings, we develop a concise notation; We say that (a, b, c, d) with respect to the preferences of the men below, indicates that man i is matched with his $P_i(x)$, where x is the i^{th} entry of this tuple. We create a similar notation for women. For example, if we say that men received an outcome $(1, 1, 2, 2)$, it means that the four men were matched with their first, first, second and second preferences, respectively, by the matching algorithm.

Figure 4.9: An example to consider various stable matches

Consider what happens when wpDA and mpDA are executed on this preference profile. The final allocations based on our new notation are described in the Table 4.1.

Both of these matches are stable. Notice that the stable match corresponding to mpDA is more favourable to the men in general, in the sense that all of them match with a woman higher up in their preference order, as compared to wpDA. The proposing party seems to obtain a more favourable outcome.

Algorithm	**Men's Matching**	**Women's Matching**
mpDA	(1,1,1,1)	(4,4,3,3)
wpDA	(3,4,3,3)	(1,1,1,1)

Table 4.1: Matching results for mpDA and wpDA algorithms

However there are more stable matches other than these two, which are not obtainable by execution of DA on this preference profile. Consider $[(1,1,2,2),(3,3,3,3)]$, which is also a stable match. This is more favourable to the men than the wpDA outcome and less favourable to the men than the mpDA outcome. Enumerating all the possible stable matches allows us to order them in terms of how favourable they are to the men/women. This leads to a lattice structure as follows:

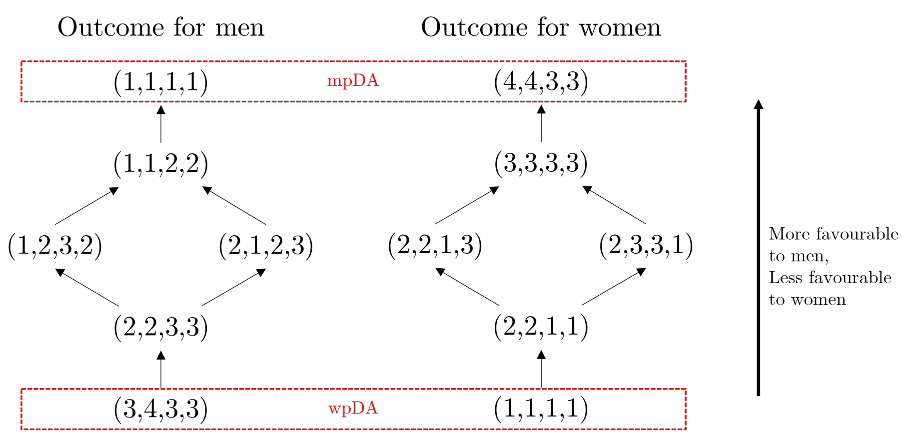

Figure 4.10: The lattice of stable matchings for the example in figure 2

The existence of such a lattice is not specific to this example but exists for any preference profile. The extremities of this lattice can be obtained by execution of DA.

With this in mind, two theorems are stated below:

Theorem 4.14 ***[Identical preference over stable matchings]***
There exists a unique stable matching that all men find at least as good as any other stable matching and there also exists a unique stable matching that they find at least as bad as the others.

Similar version of the above theorem extends to women.

Theorem 4.15 ***[Reversed preferences for men and women]***
For any two distinct stable matchings Q and R, if all men find Q at least as favourable as R, then all women find R at least as favourable as Q.

References

[GSDA] D. Gale and L. S. Shapley, "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962. Accessed: Jan. 17, 2025.