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

Lecture 6: Manipulation in Stable matching

Lecturer: Swaprava Nath *Scribe(s): Nivesh Aggarwal*

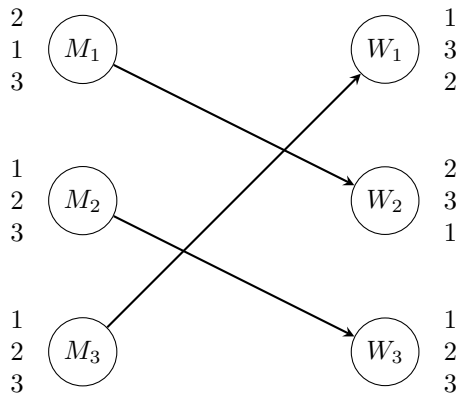
Disclaimer: *These notes aggregate content from several texts and have not been subjected to the usual scrutiny deserved by formal publications. If you find errors, please bring to the notice of the Instructor.*

6.1 Manipulation

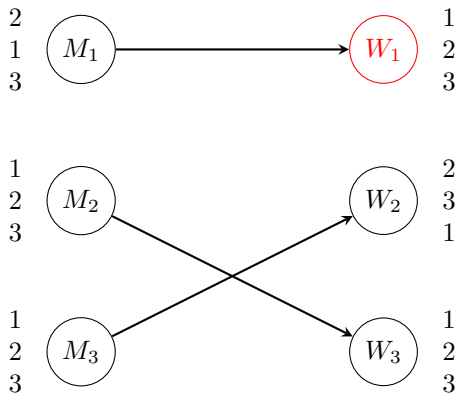
Definition 6.1 (Strategy Proof) *A matching X is manipulable if $\exists i \in N = M \cup W$ and a profile P s.t. $X_i(P_i', P_{-i})$ P_i $X_i(P_i, P_{-i})$. A matching algorithm is strategy proof if it is not manipulable at **any profile** by any agent.*

6.1.1 Manipulation by non-proposers

DA is manipulable. In men proposing DA it is not manipulable by men but it is manipulable by women (non-proposers) as shown in Figure 6.1



(a) Matching (X) returned by men-proposing DA algorithm on profile P



(b) Matching (Y) returned by men-proposing DA algorithm on profile P' when W_1 misreports

Figure 6.1: Manipulation by women (non-proposers) in DA

6.2 Optimal Manipulation and Stability

Definition 6.2 (Optimal Manipulation) *Among all possible manipulations a non-proposing agent can perform , the one that gives her/his the best possible match is the optimal manipulation by that agent.*

Definition 6.3 (Single elevation) *Raising the rank of a single man in the preference profile.*

Theorem 6.4 *Any optimal manipulation for a woman in a men-proposing DA can also be achieved via a "single elevation" misreport. Therefore finding the optimal misreport is $O(n^2)$*

Theorem 6.5 *The men-proposing DA outcome after optimal manipulation by a woman is stable w.r.t the original profile.*

Observation - In a single elevation misreport(SEM) of W_i if M' proposed to W_i in the original matching profile the M' also proposes to it in the new profile P' if W_i elevated M'.

Proof: $X = mp - DA(P_W, P_{-W})$ be the outcome of the men-proposing DA on the preference profiles P_W, P_{-W} and $X' = mp - DA(P'_W, P_{-W})$ be the outcome of the mp-DA on the manipulated profile P'_W, P_{-W} where P'_W is the optimal manipulation for W.

We prove by contradiction. Suppose X' is not stable.

Case 1: The blocking pair is M, W' where $W' \neq W$.

Since we did not change $P_{W'}$ or P_M and they are blocking pairs in P they must be a blocking pair in P' this would imply that X' is not stable w.r.t. P' which is not possible.

Case 2: The blocking pair is M, W.

We consider a new profile P'' created by elevating only M to the top. M must propose to W in P' as it has gotten a lower preference and hence must have proposed to W and would have been rejected. Hence by our observation M must propose to W in P'' and W will accept. But this would imply there exists a profile P'' which gives a better outcome than P'.

But this implies that P' is not the optimal manipulation which is a **contradiction**. Hence X' is stable ■

Theorem 6.6 *There does not exist a deterministic stable matching algorithm that is strategy-proof.*

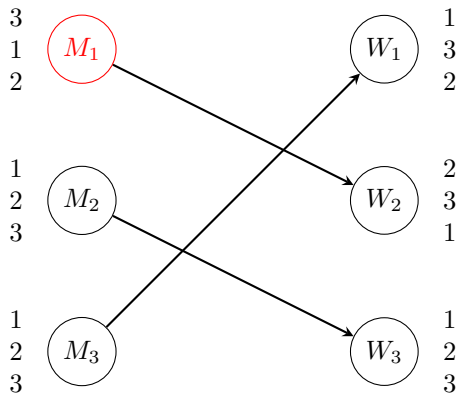


Figure 6.2: Matching (X') returned by manipulating the profile P'' of M_1

In profile P the only stable matchings that exist are X and Y. We know that X is stable as we get it from mp-DA and Y is stable because we get it from mp-DA in the SEM profile P'.

We can also see that these are the only 2 stable matching possible as if M_1 is matched with W_3 then M_1 and W_1 are always a blocking pair. Similarly M_3 can't be matched to W_3 as then M_3 and W_2 is always a blocking pair. Hence M_2 must match to W_3 and only 2 stable matchings are possible.

In profile P' (Figure 6.1b) M_1 and W_1 must be matched as both are each other's top preference. Also as before M_3 can't match to W_3 so only 1 stable matching is possible. In profile P" M_3 and W_1 must be matched as both are each other's top preference (Figure 6.2).

Now only 1 stable matching is possible as M_1 and W_2 prefer each other over the other available options (Figures 6.1a and 6.2). Hence any deterministic algorithm will return either X or Y. If it returns X then W_1 can manipulate and if it gives Y M_1 can manipulate.