

## Lecture 3

# Manipulation in Stable Matching Problem

# Stable Matching Problem

$w_1 > w_2 > w_3$



$m_3 > m_2 > m_1$

$w_2 > w_1 > w_3$



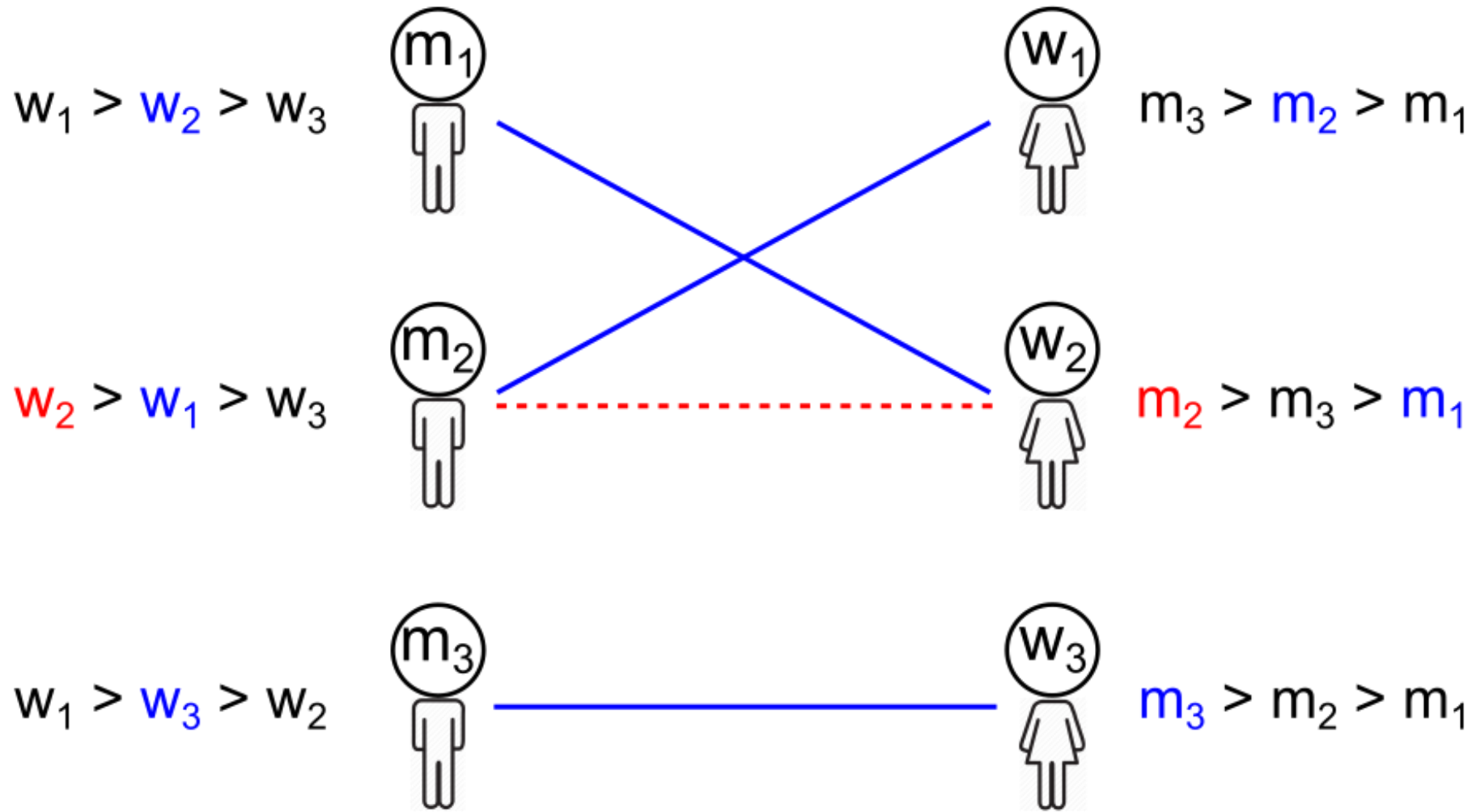
$m_2 > m_3 > m_1$

$w_1 > w_3 > w_2$



$m_3 > m_2 > m_1$

# Stable Matching Problem



A matching is **stable** if there is no **blocking pair**.



## COLLEGE ADMISSIONS AND THE STABILITY OF MARRIAGE

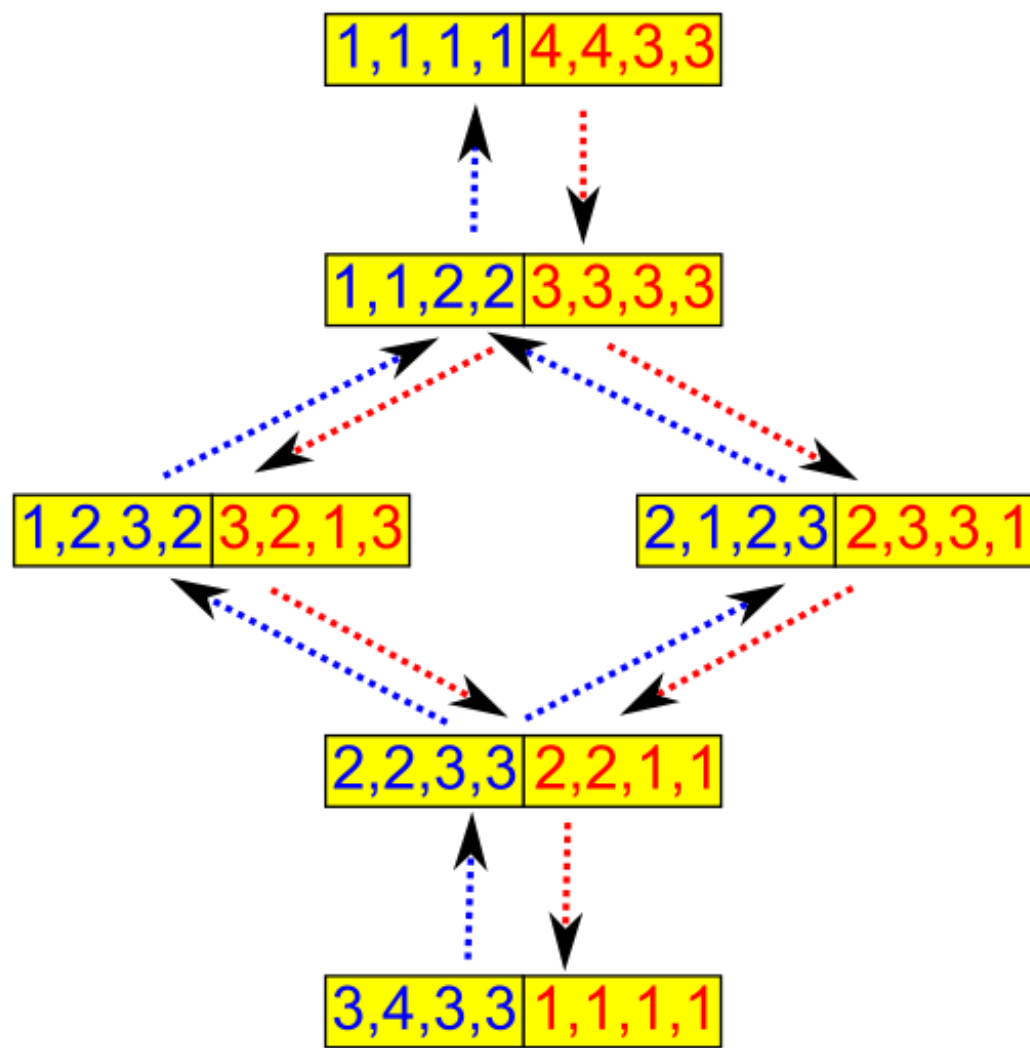
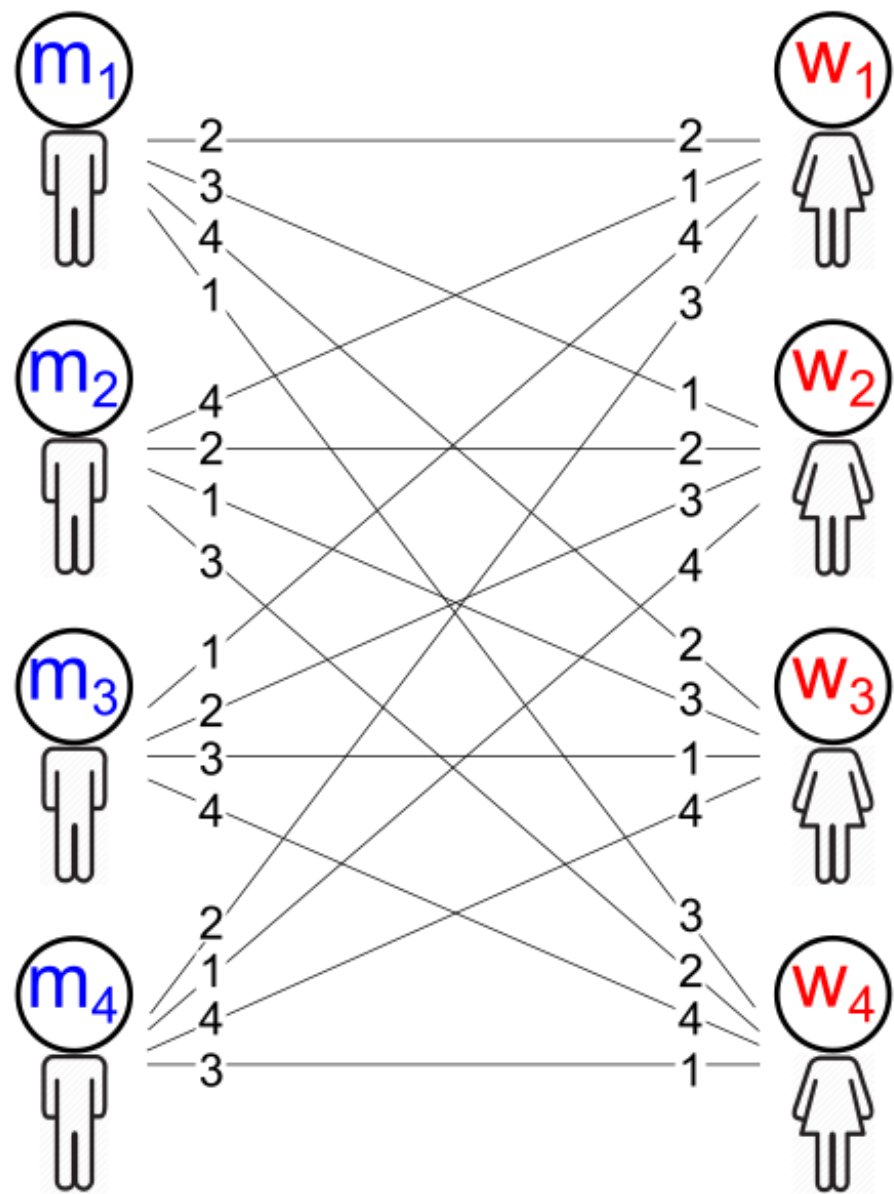
D. GALE\* AND L. S. SHAPLEY, Brown University and the RAND Corporation

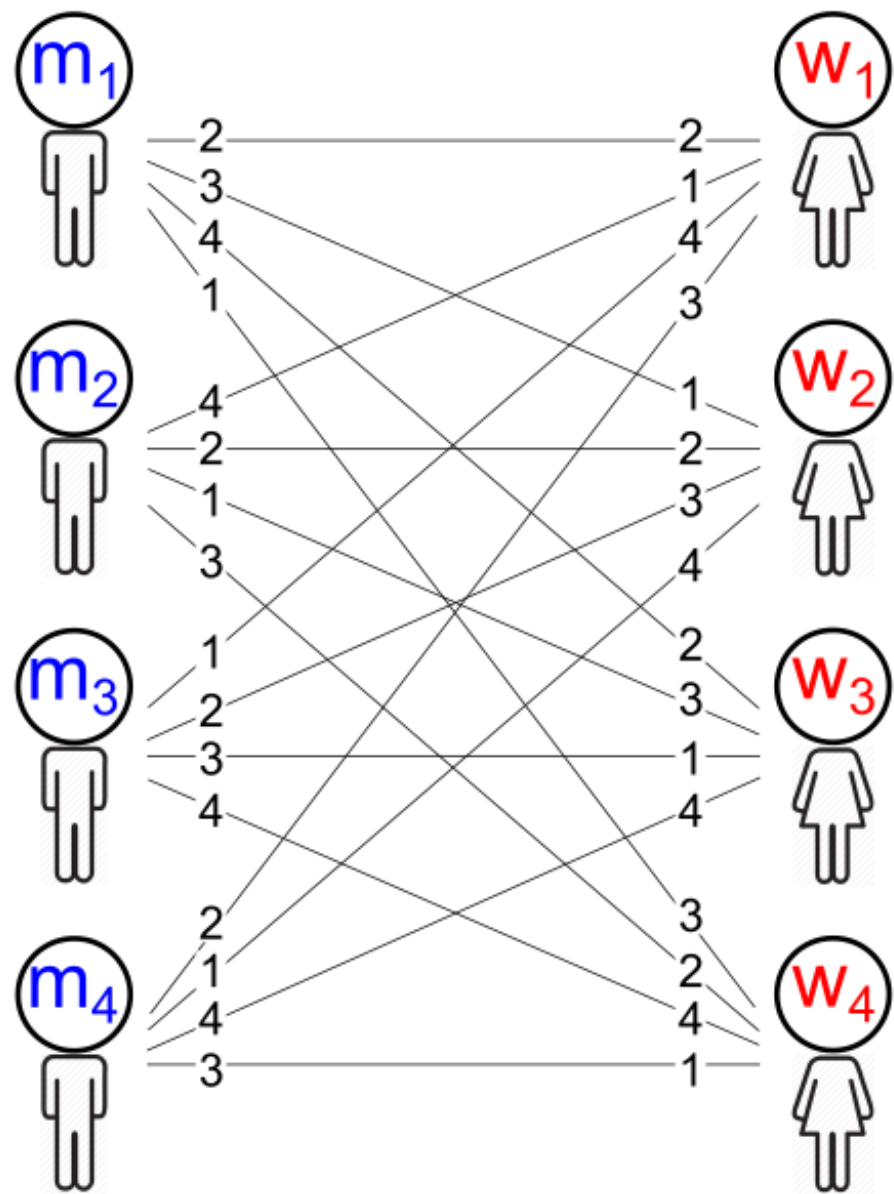


Source: *The American Mathematical Monthly*, Jan., 1962, Vol. 69, No. 1 (Jan., 1962), pp. 9-15

Given any preference profile, a stable matching for that profile always exists and can be computed in polynomial time.



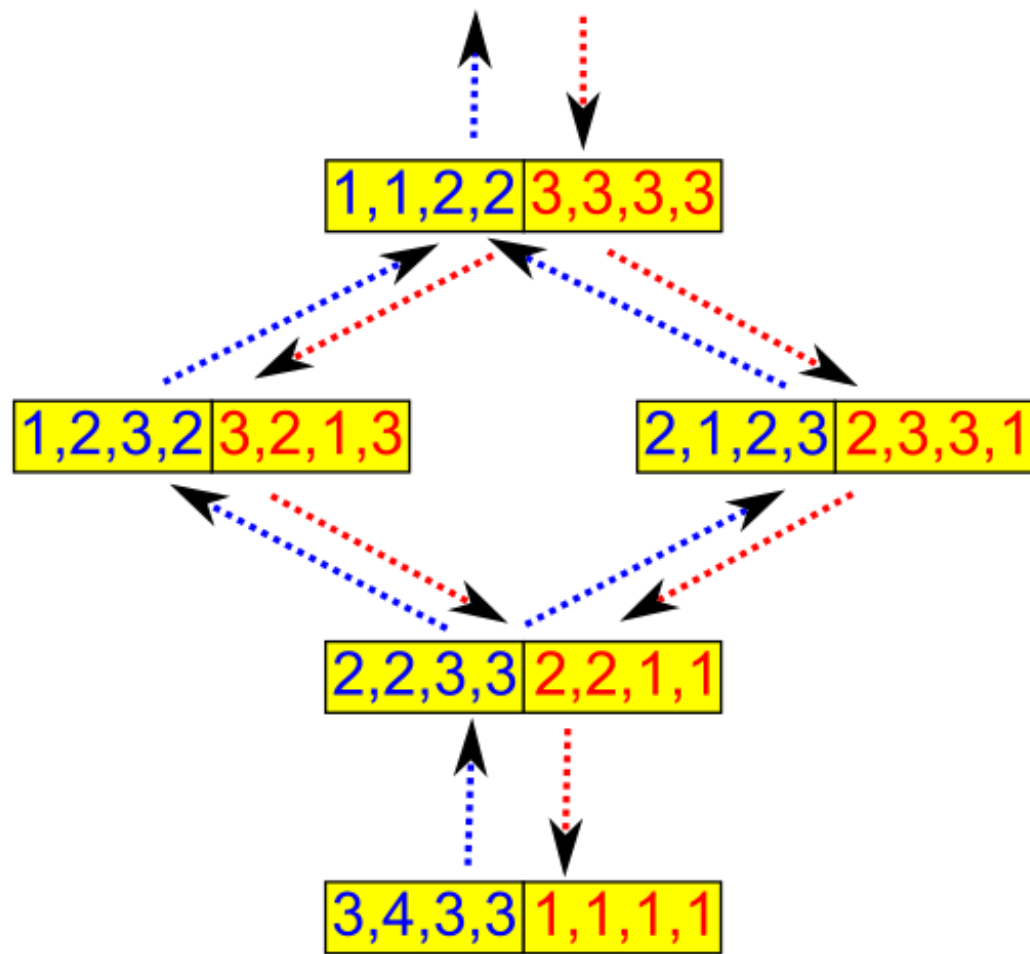


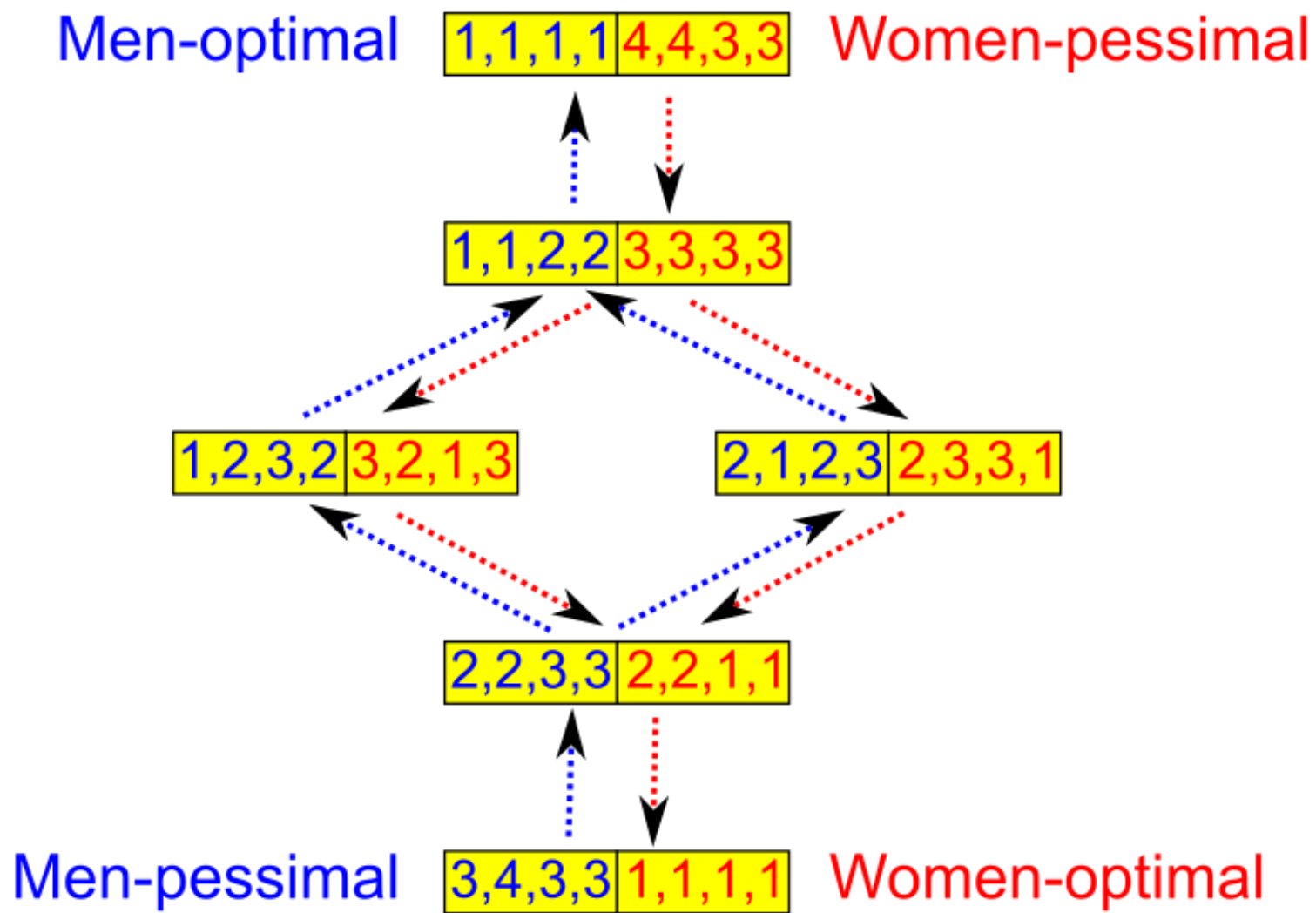
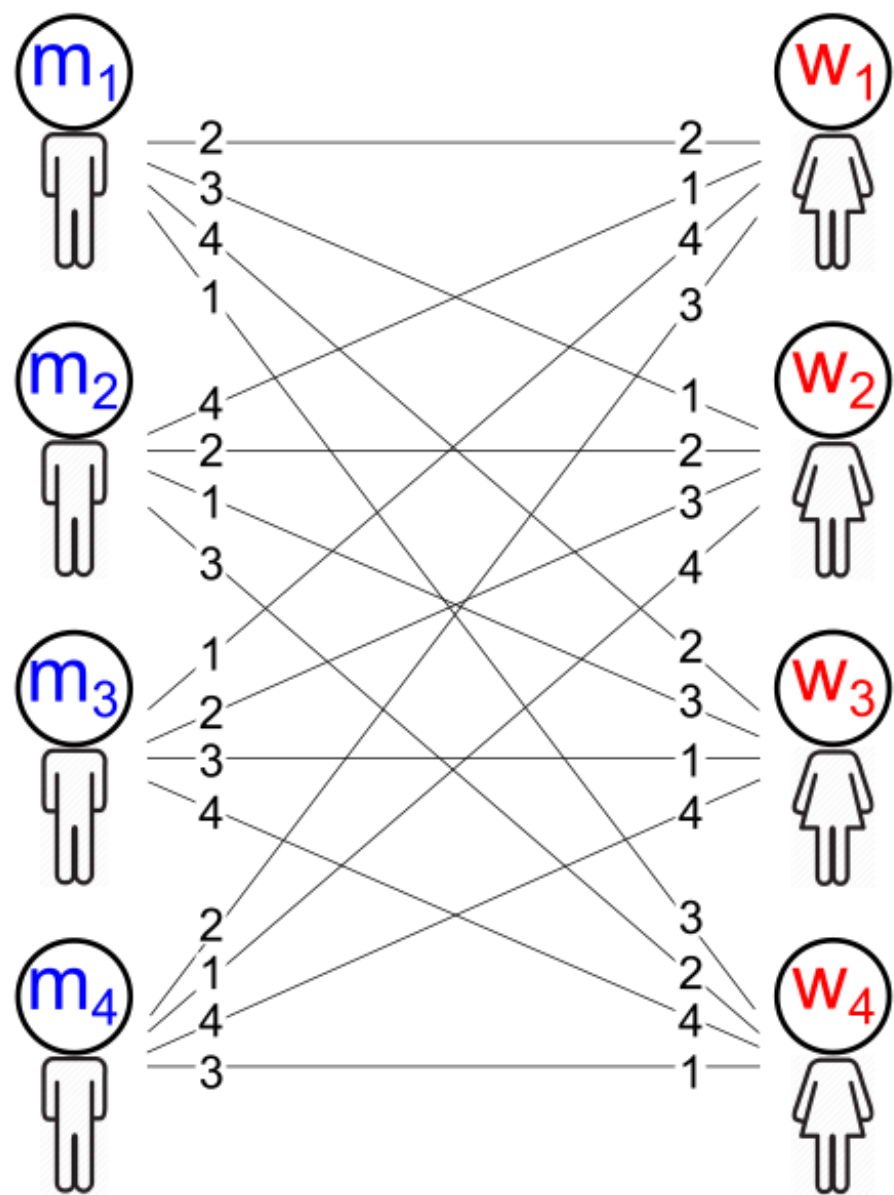


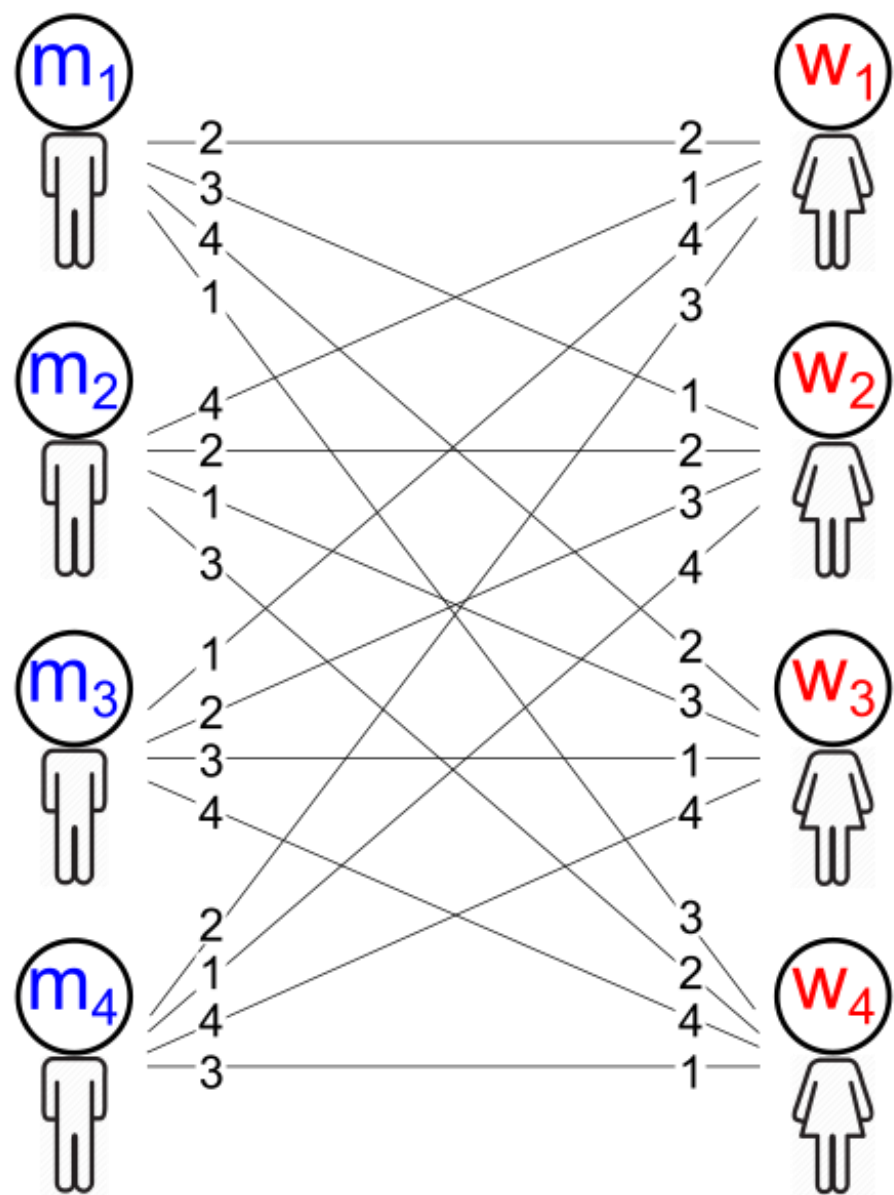
Men-optimal

1,1,1,1 | 4,4,3,3

Women-pessimal

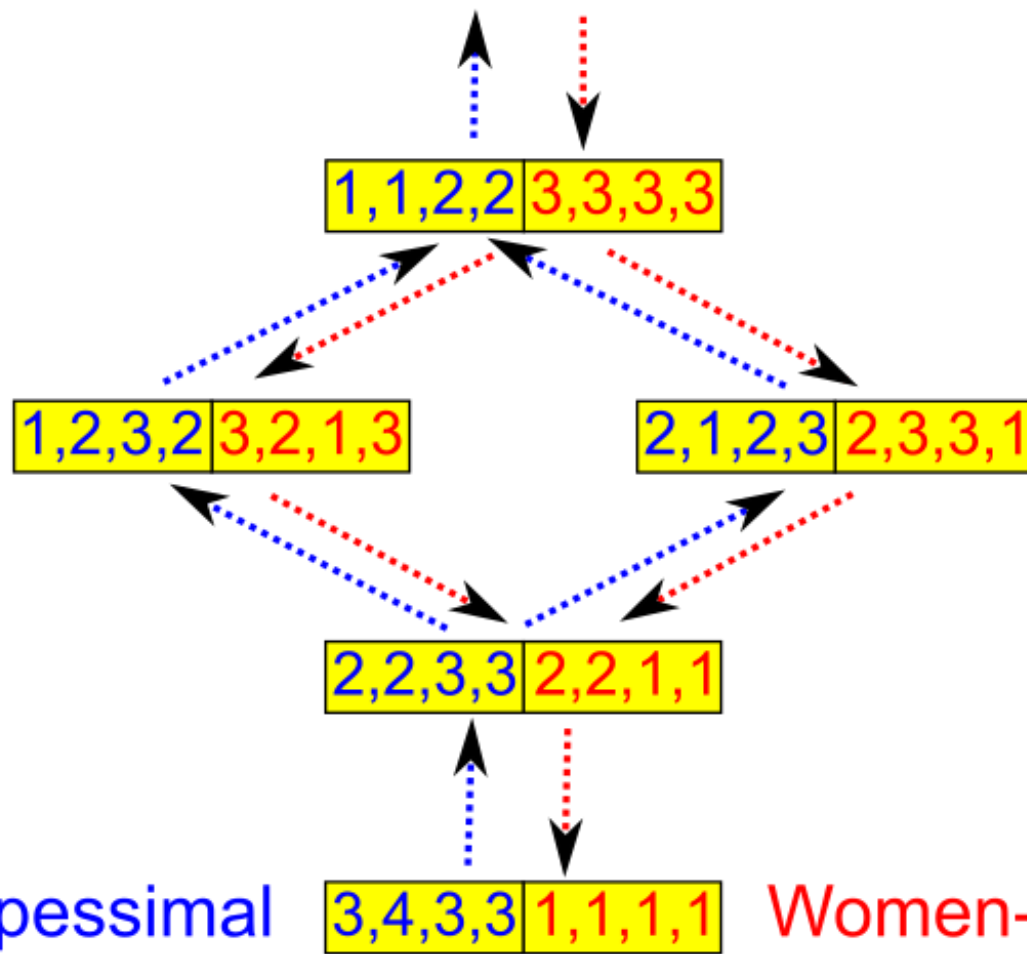




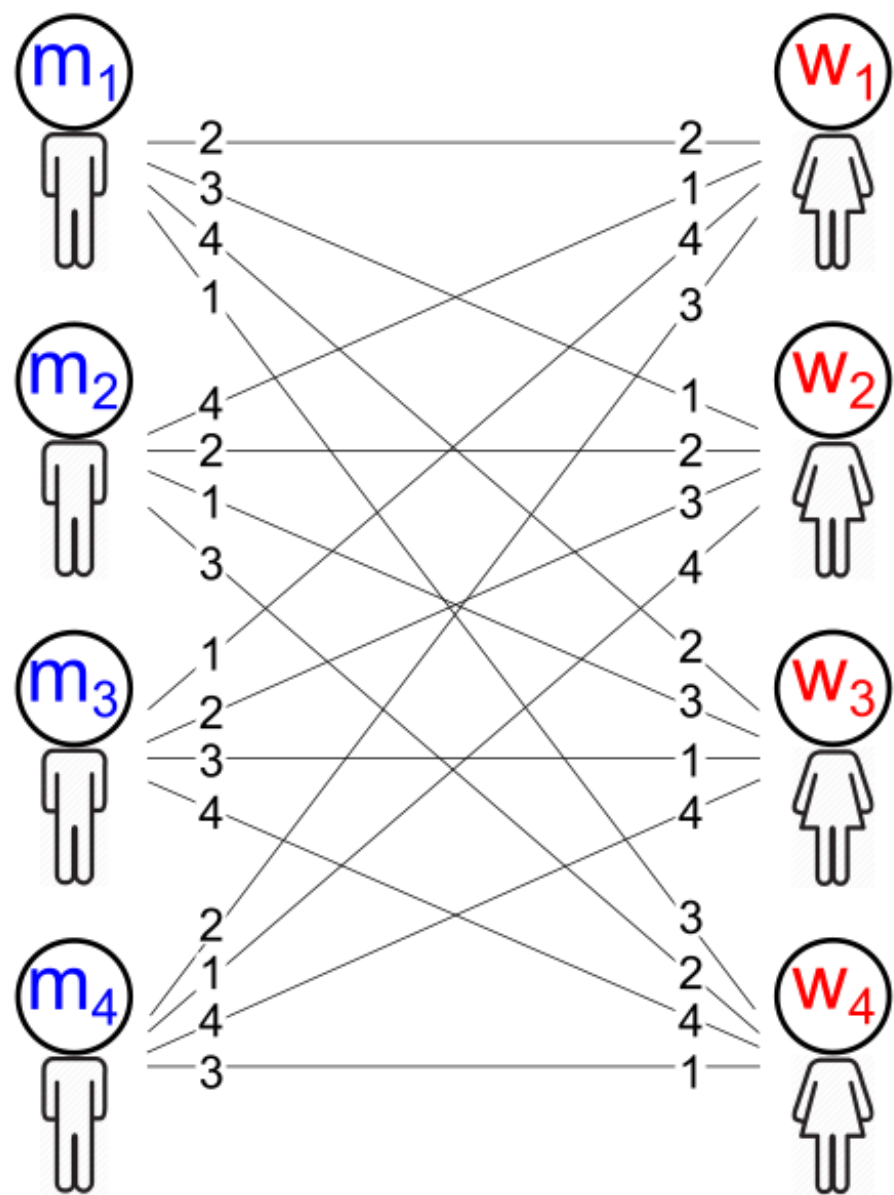


Men-proposing DA algorithm computes this

Men-optimal  $1,1,1,1 \mid 4,4,3,3$  Women-pessimal

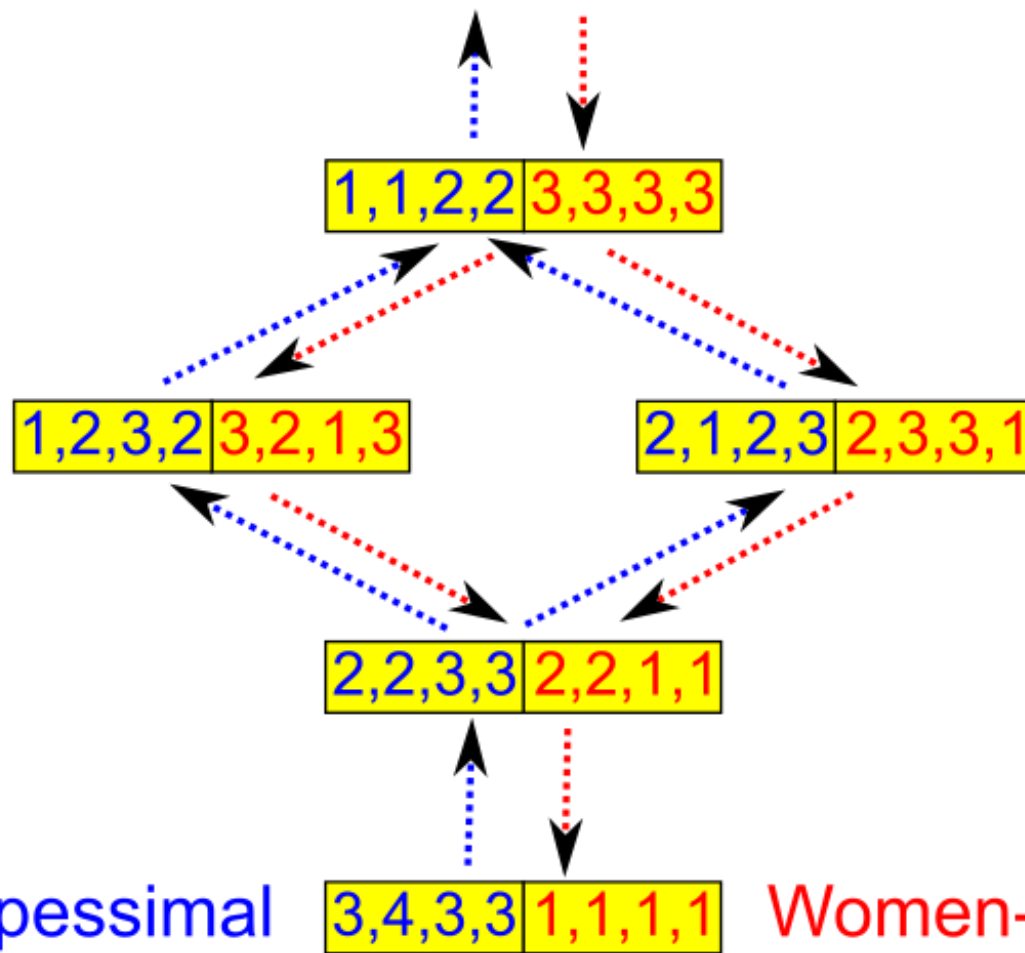


Men-pessimal  $3,4,3,3 \mid 1,1,1,1$  Women-optimal



Men-proposing DA algorithm computes this

Men-optimal  $1, 1, 1, 1 \mid 4, 4, 3, 3$  Women-pessimal

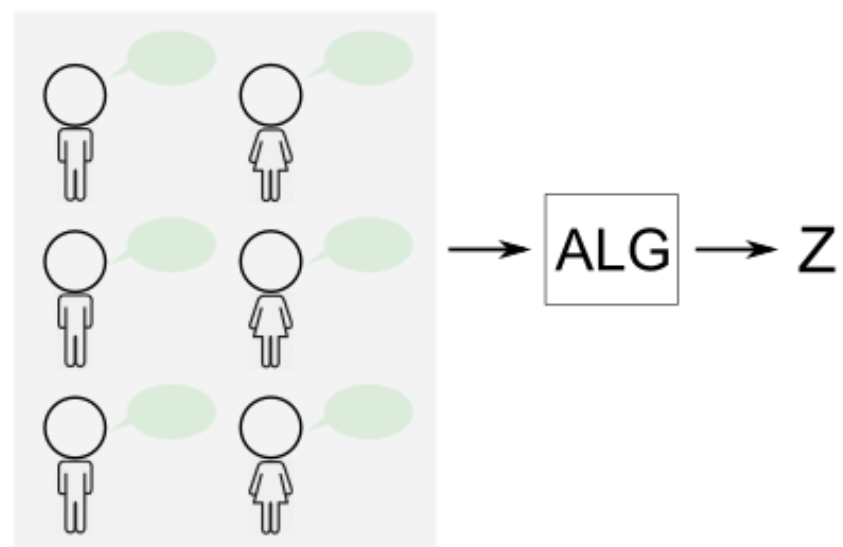
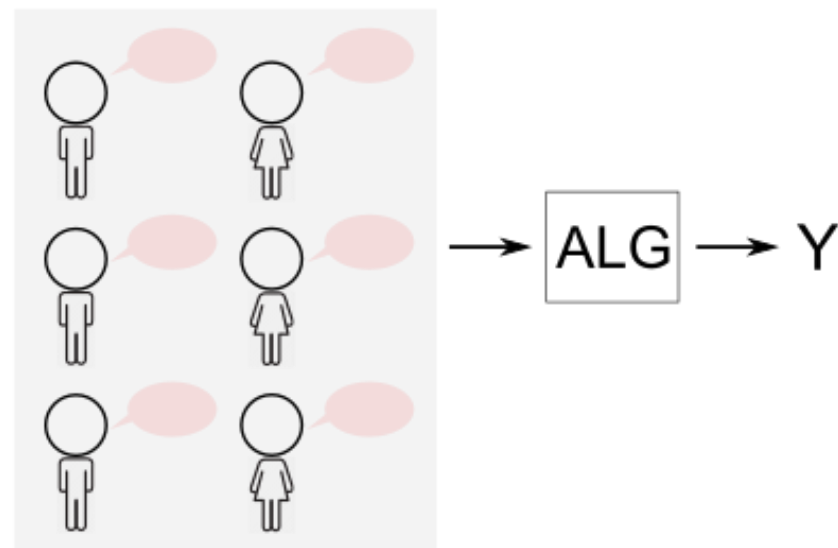
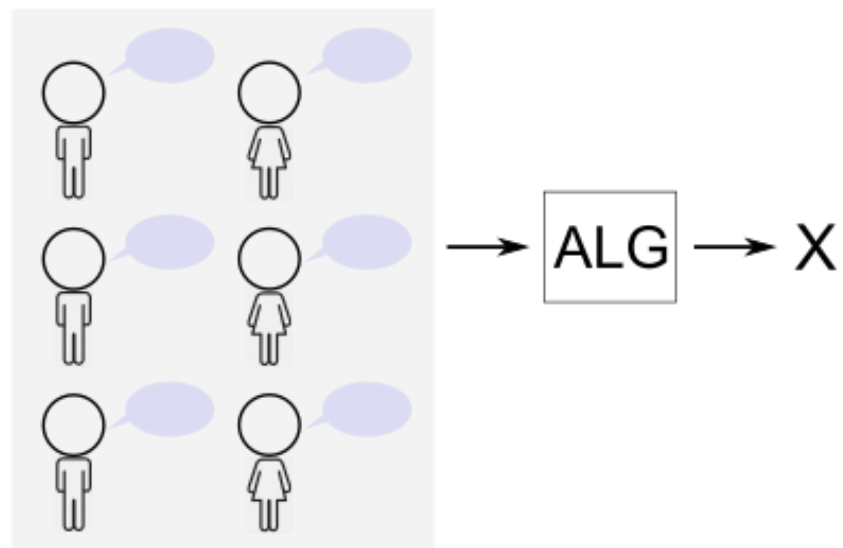


Women-proposing DA algorithm computes this

Can an agent get a better partner under DA algorithm  
by misreporting his/her preferences?

# Strategyproofness

# Strategyproofness



...



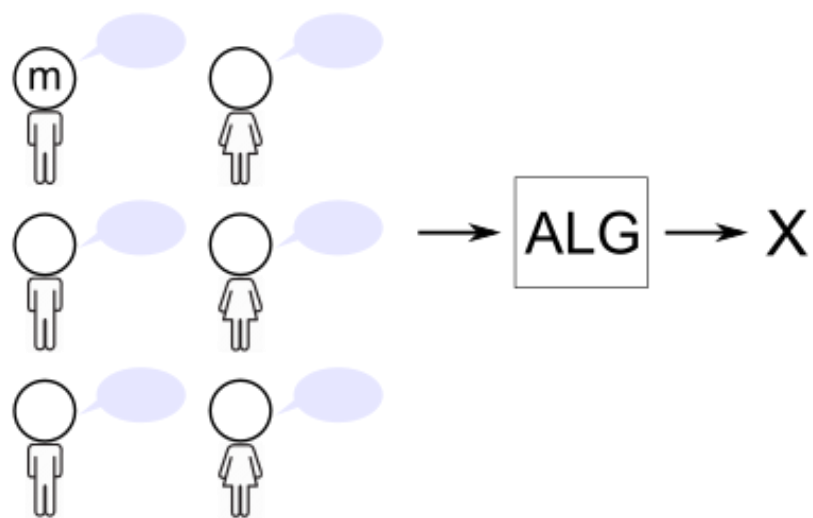
# Strategyproofness

An algorithm fails strategyproofness if, for some set of input preferences, some agent can unilaterally misreport and improve (w.r.t. true preference).

# Strategyproofness

An algorithm fails strategyproofness if, for some set of input preferences, some agent can unilaterally misreport and improve (w.r.t. true preference).

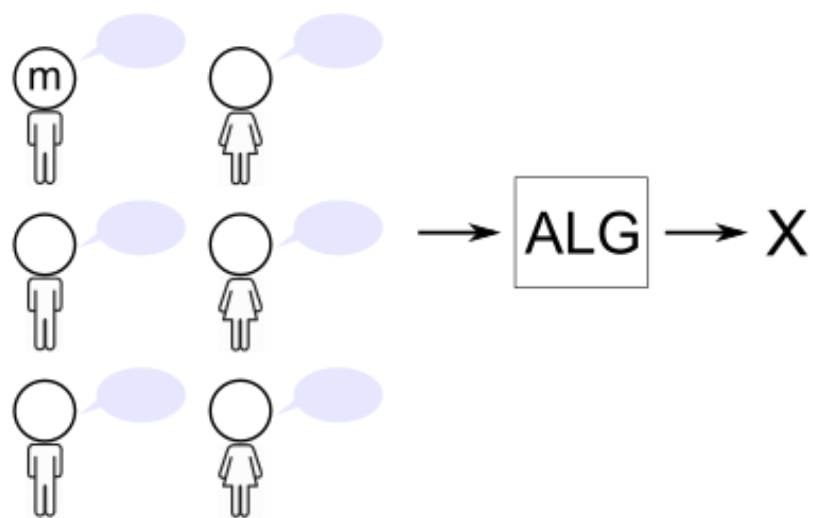
True profile  $P = (P_{-m}, P_m)$



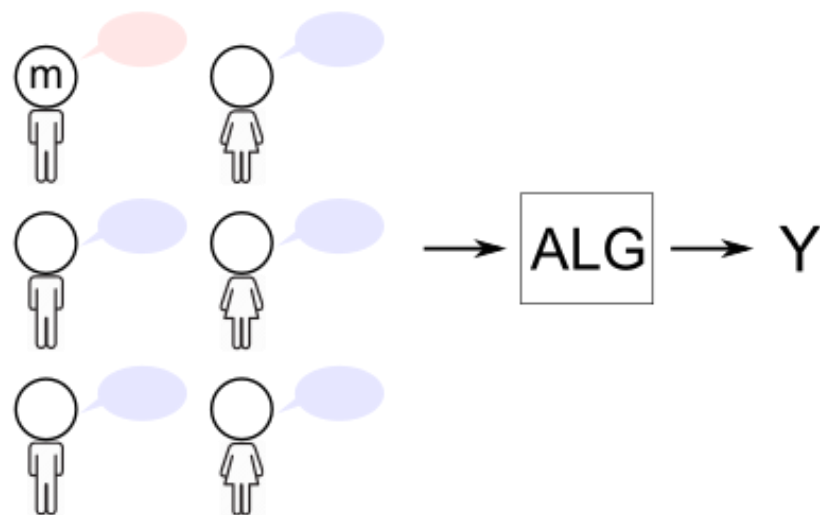
# Strategyproofness

An algorithm fails strategyproofness if, for some set of input preferences, some agent can unilaterally misreport and improve (w.r.t. true preference).

True profile  $P = (P_{-m}, P_m)$



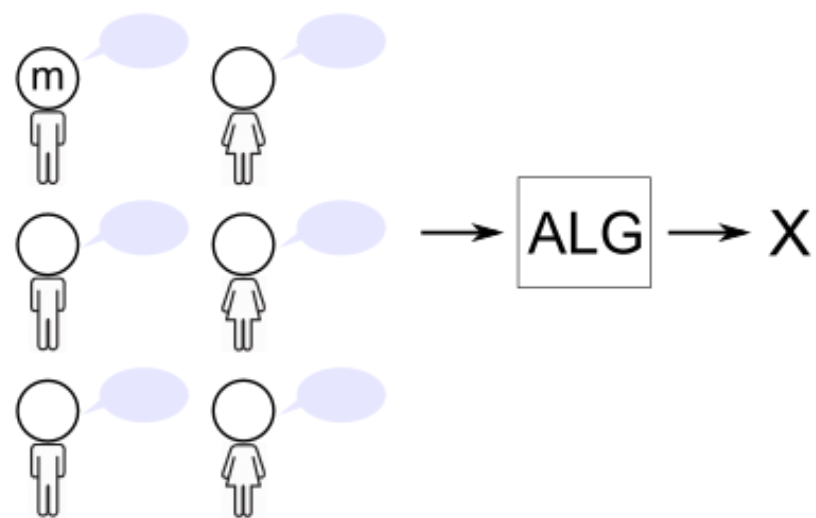
Manipulated profile  $P' = (P_{-m}, P'_m)$



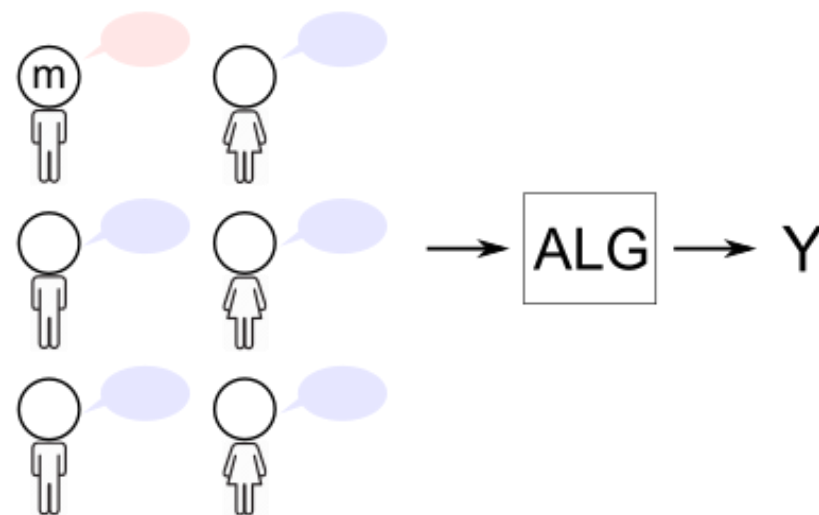
# Strategyproofness

An algorithm fails strategyproofness if, for some set of input preferences, some agent can unilaterally misreport and improve (w.r.t. true preference).

True profile  $P = (P_{-m}, P_m)$



Manipulated profile  $P' = (P_{-m}, P'_m)$



True preference list  $P_m$  ...  $> Y(m) > \dots > X(m) > \dots$

# Strategyproofness

An algorithm fails strategyproofness if, for some set of input preferences, some agent can unilaterally misreport and improve (w.r.t. true preference).

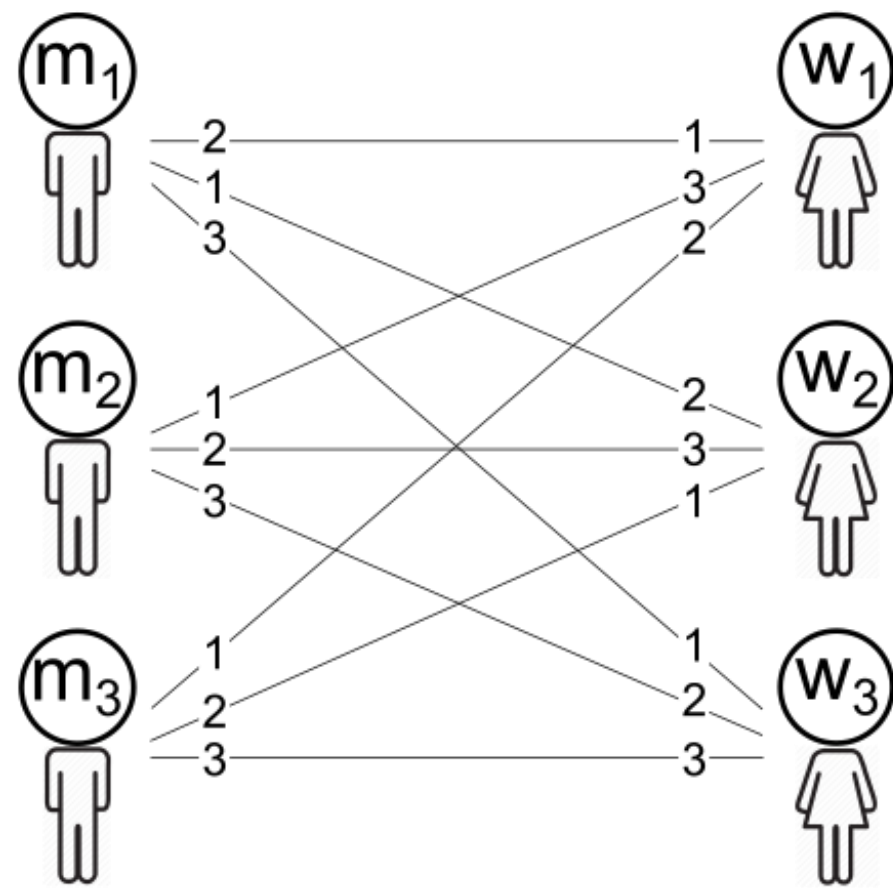
An algorithm is **strategyproof** if it does not fail strategyproofness on any input.

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is not strategyproof.

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

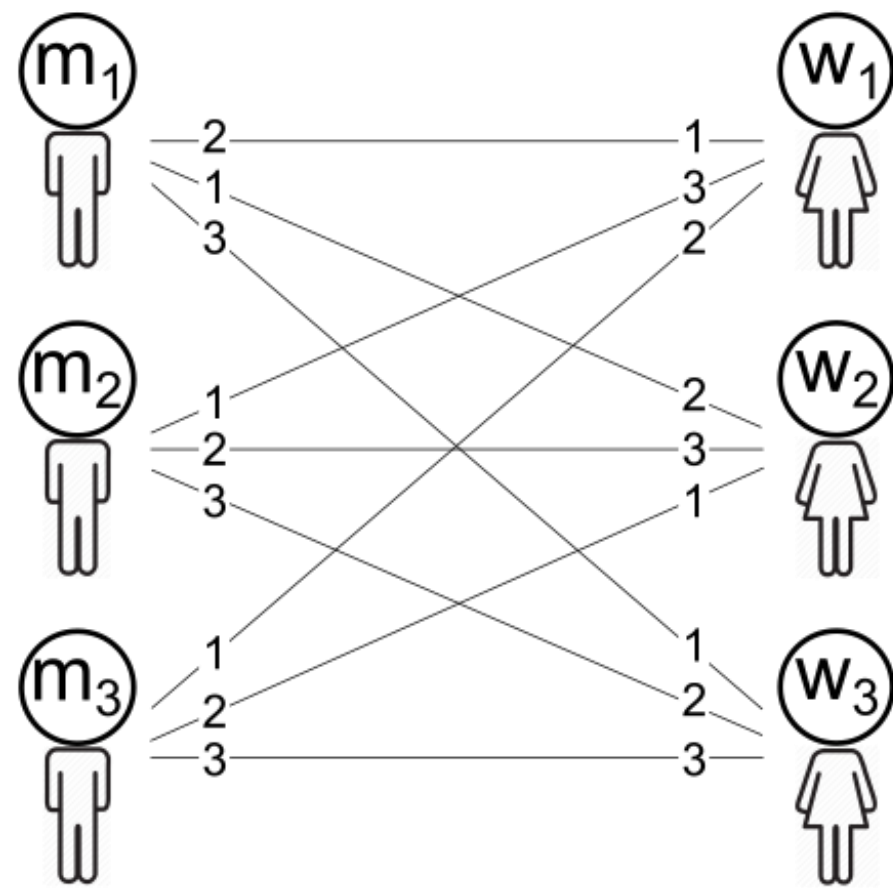
DA algorithm is not strategyproof.



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is not strategyproof.

Round 1

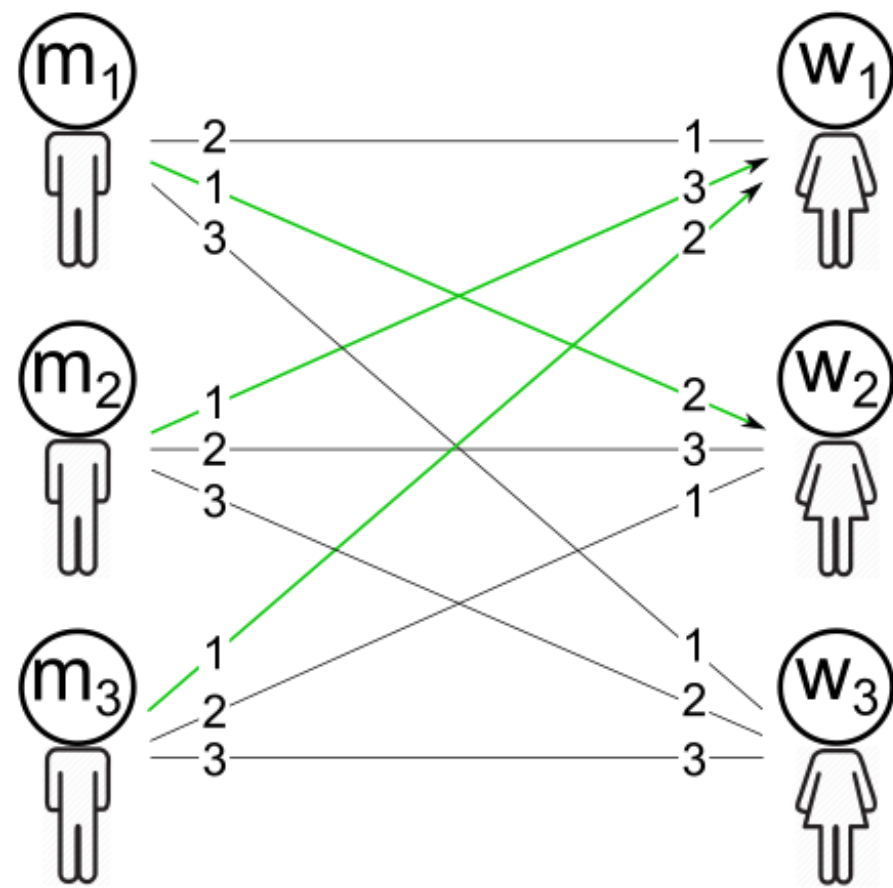




[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is not strategyproof.

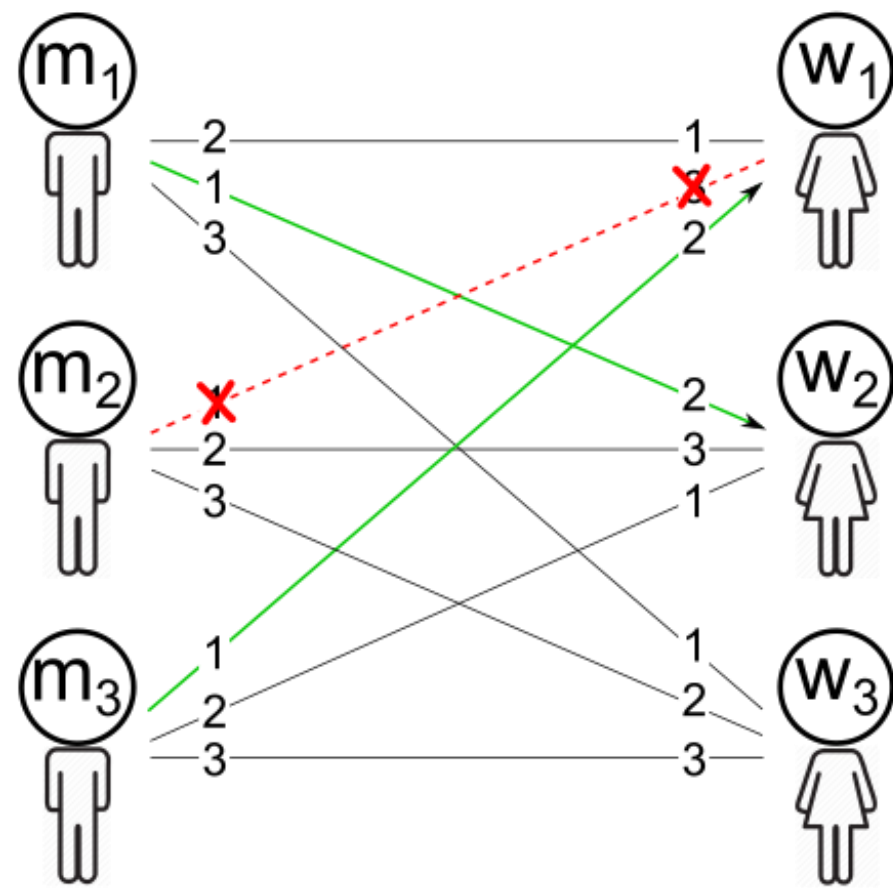
Round 1



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is not strategyproof.

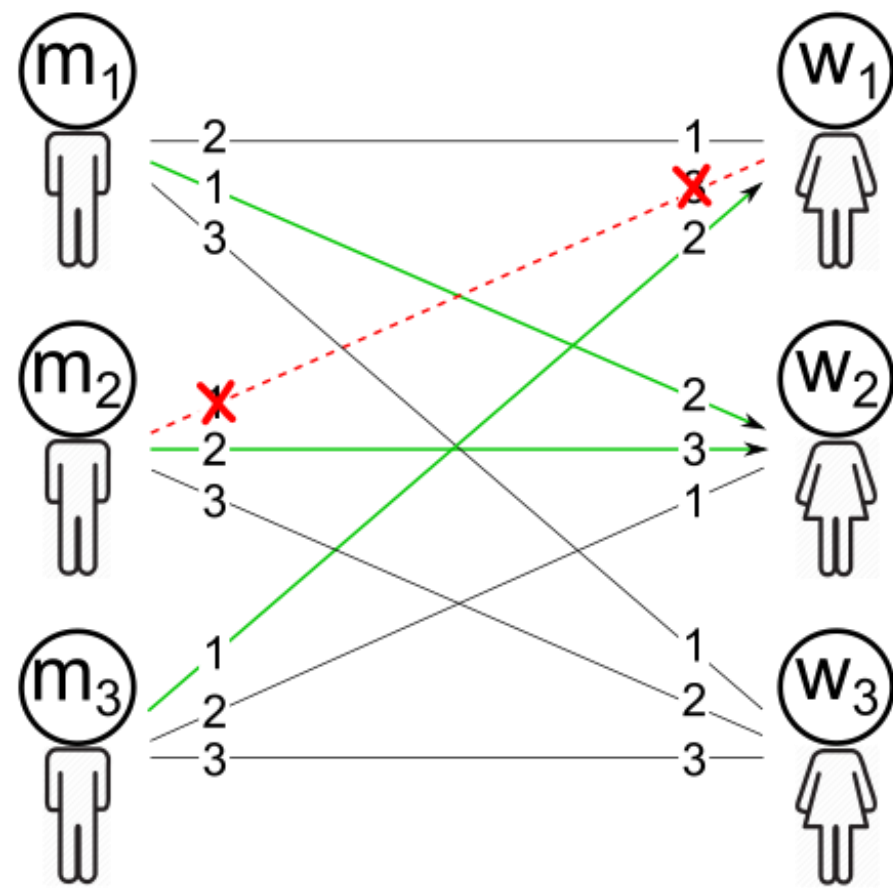
Round 1



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is not strategyproof.

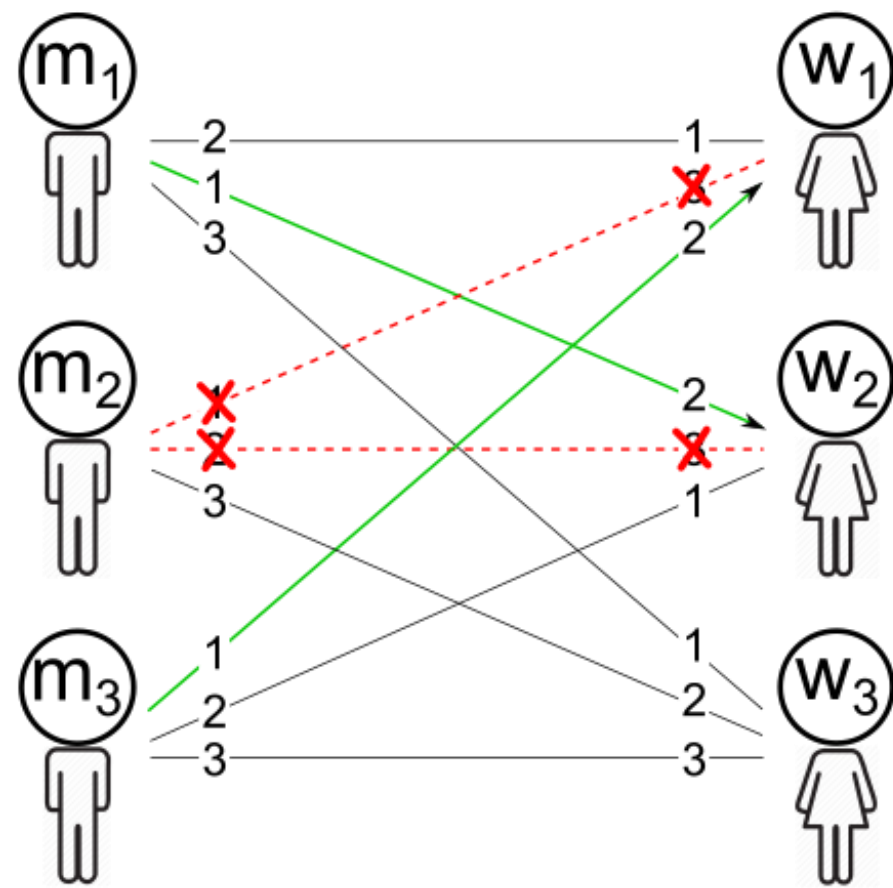
Round 2



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is not strategyproof.

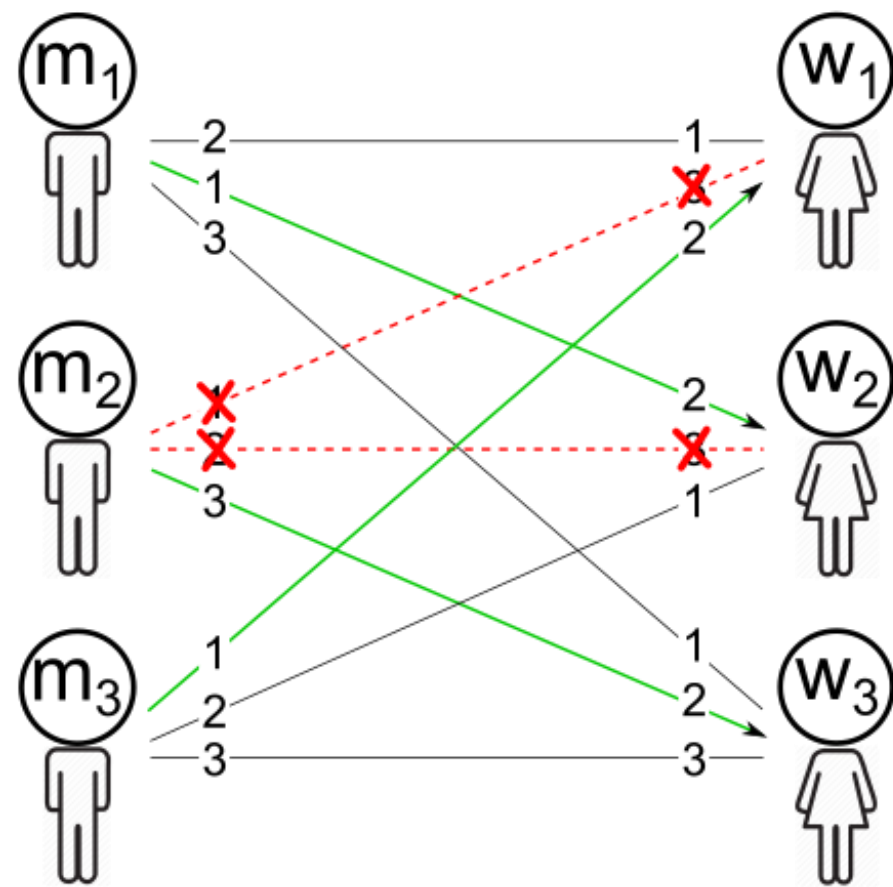
Round 2



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

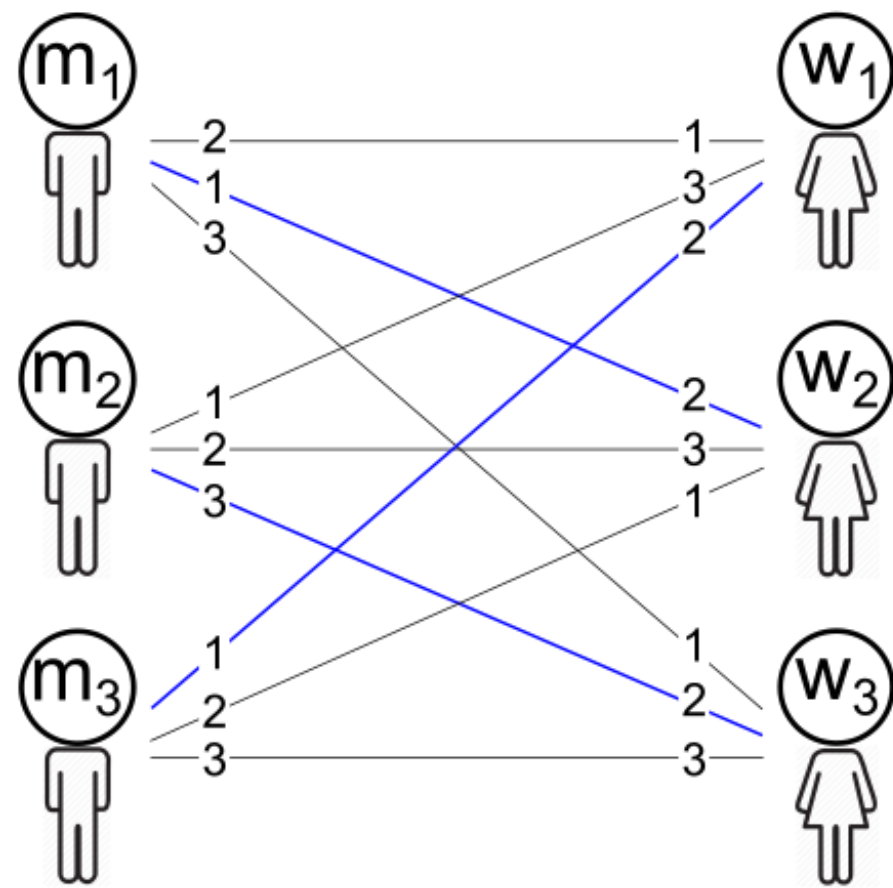
DA algorithm is not strategyproof.

Round 3



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

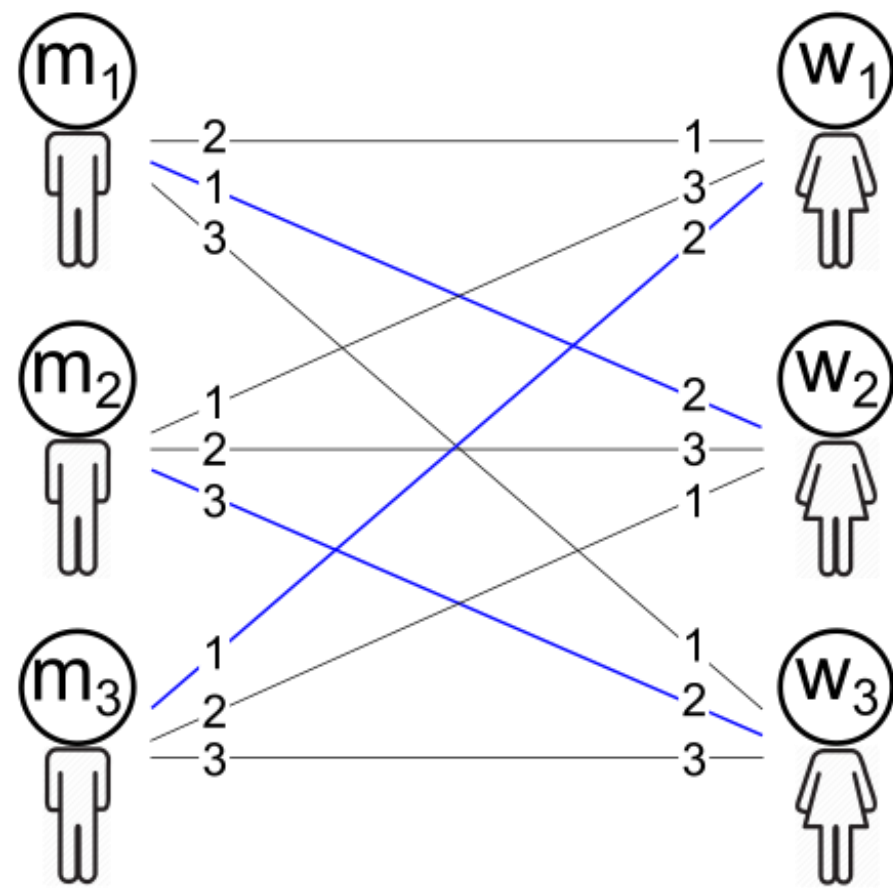
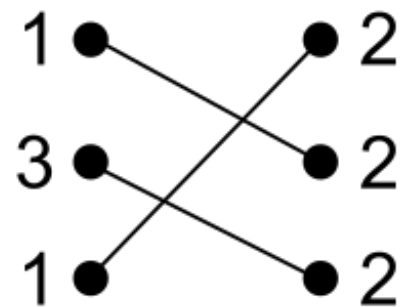
DA algorithm is not strategyproof.



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is not strategyproof.

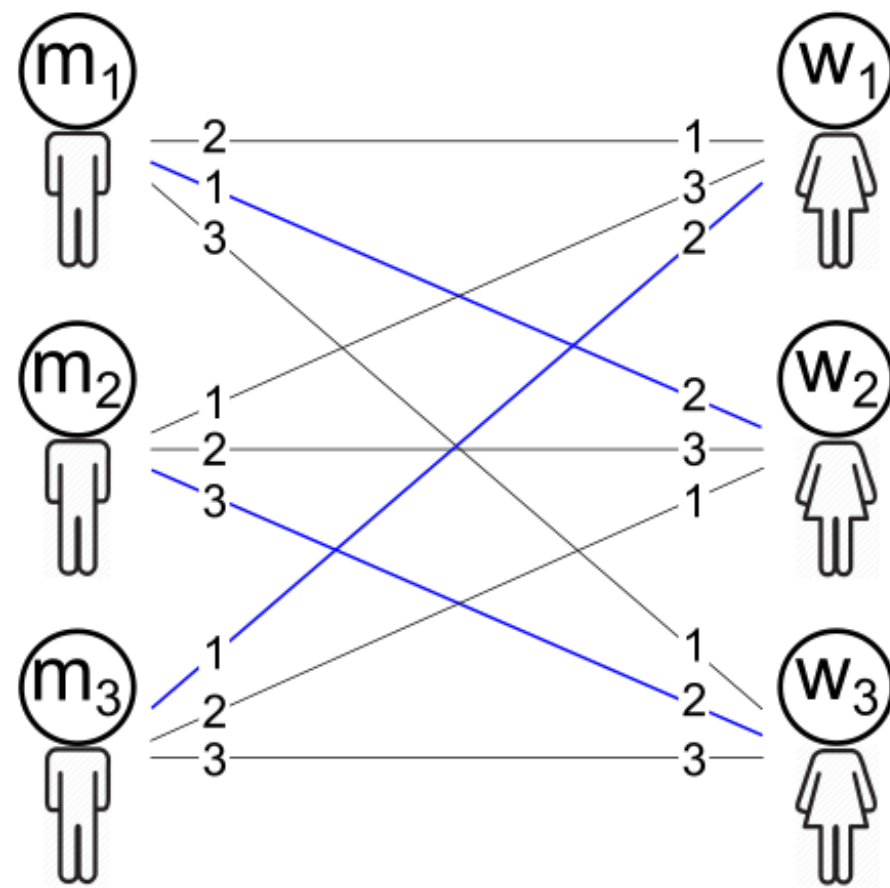
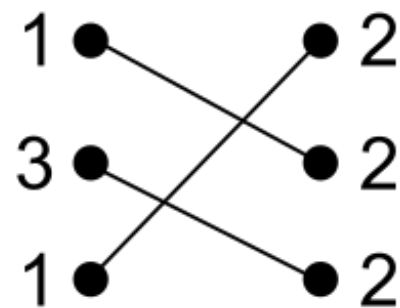
Outcome for true prefs



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs



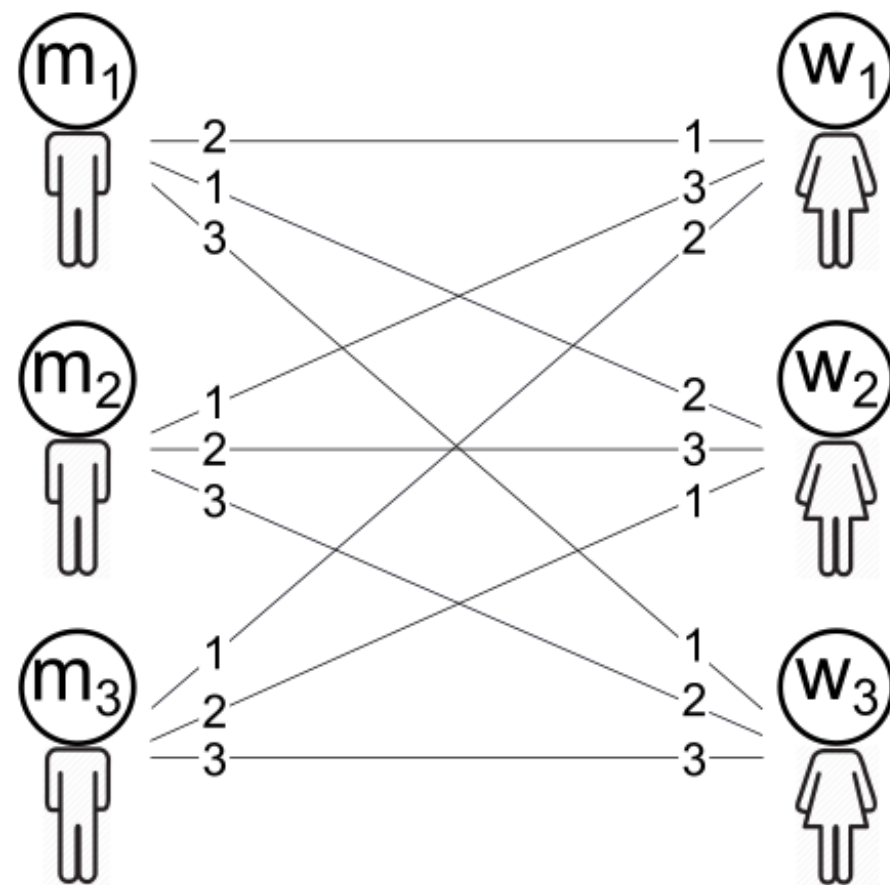
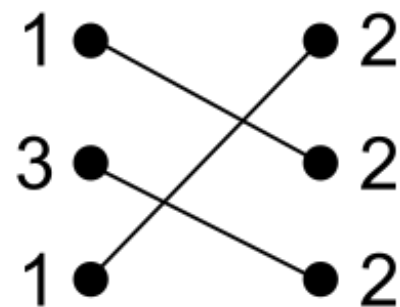
Can I get a better partner by misreporting my preferences?



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

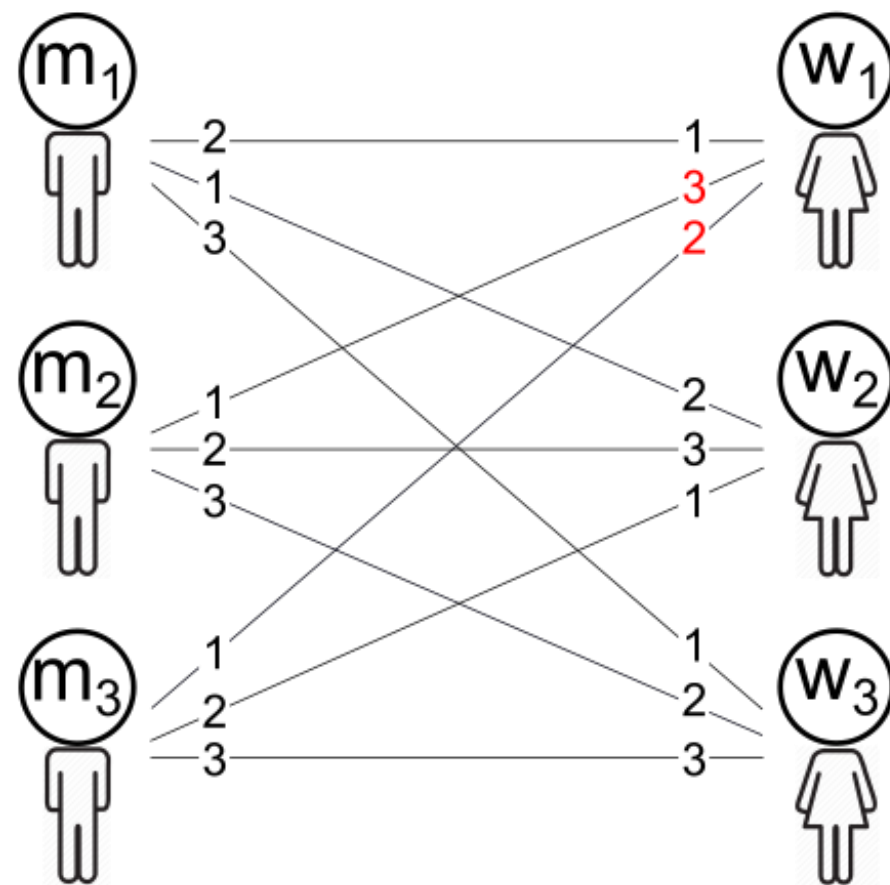
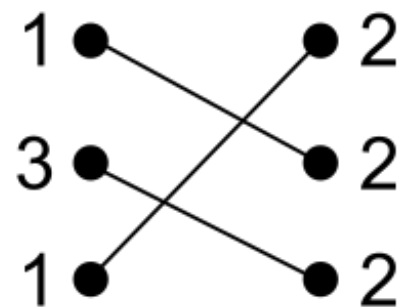


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

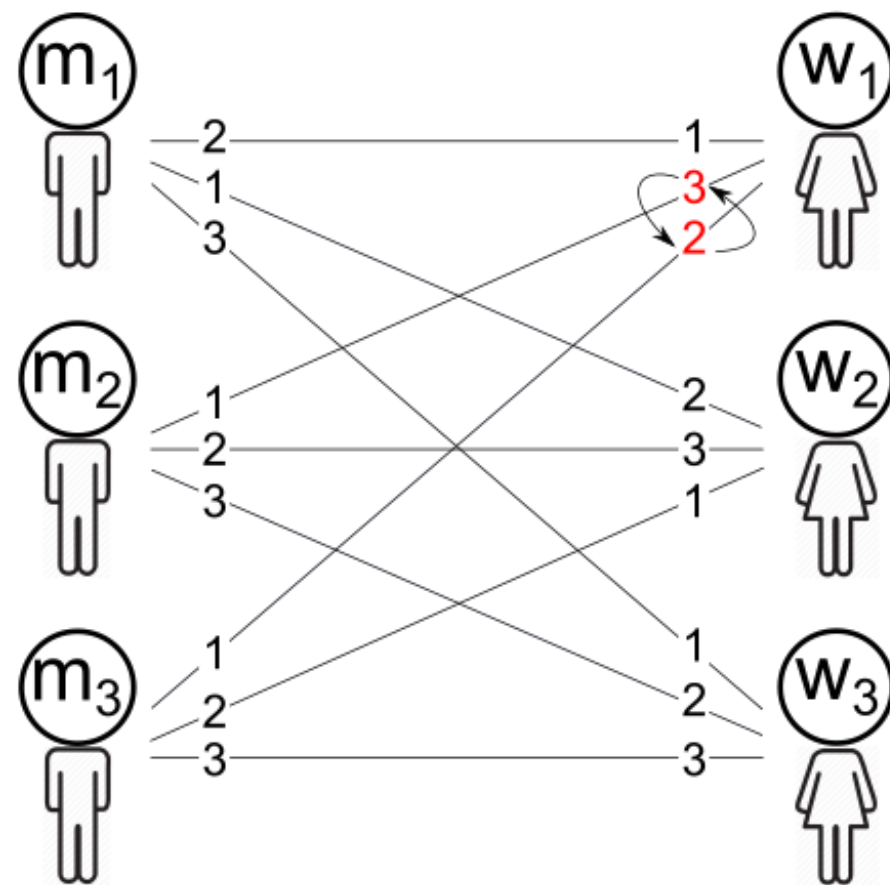
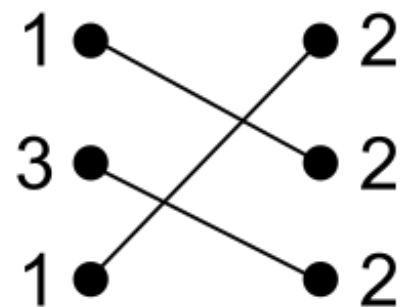


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

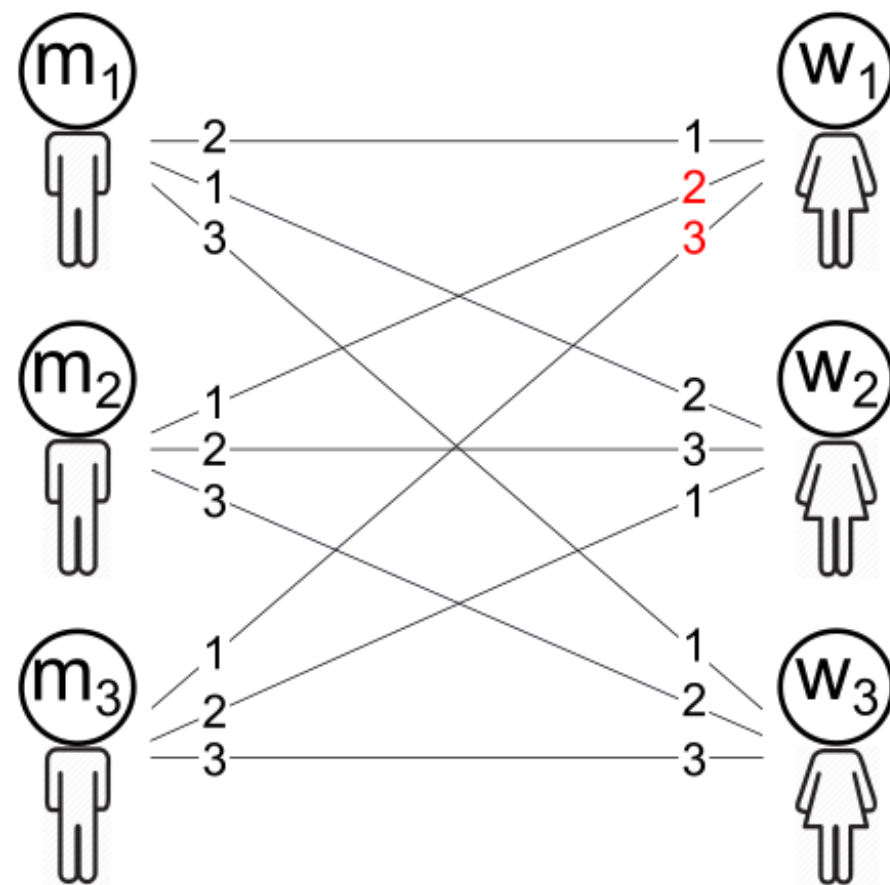
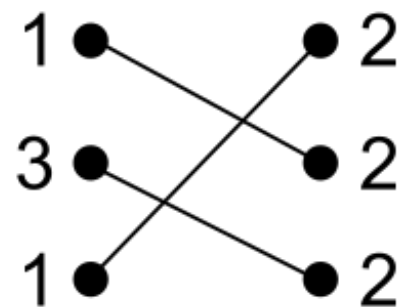


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

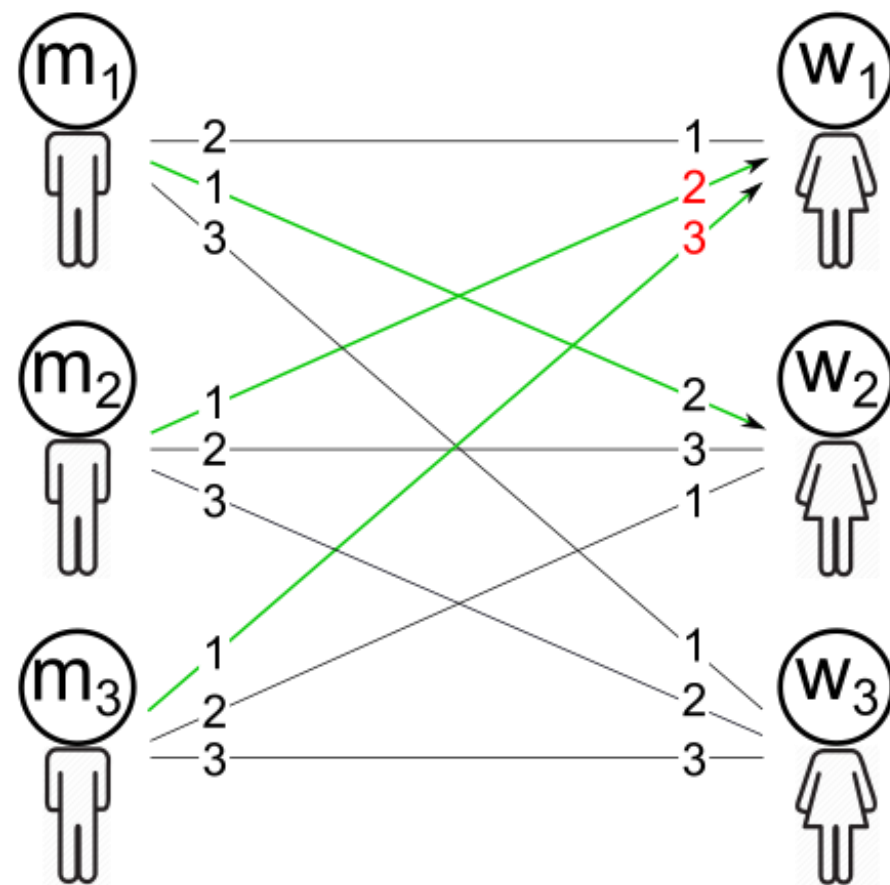
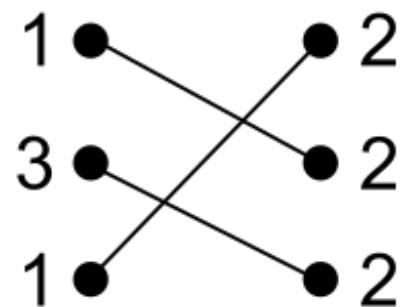


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

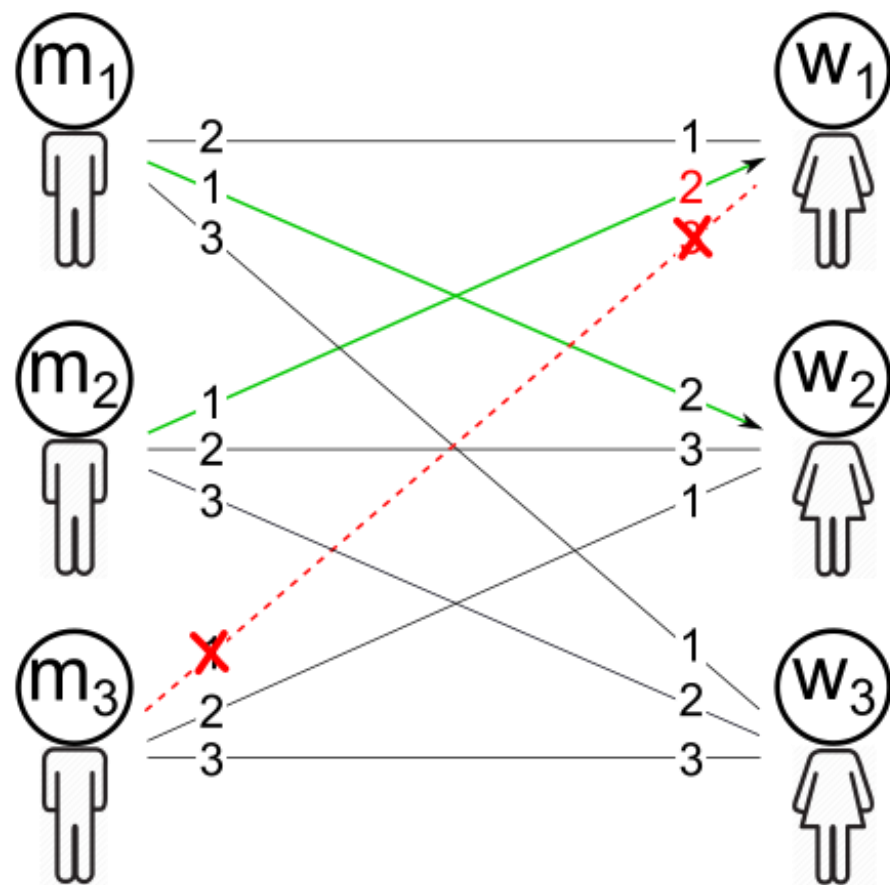
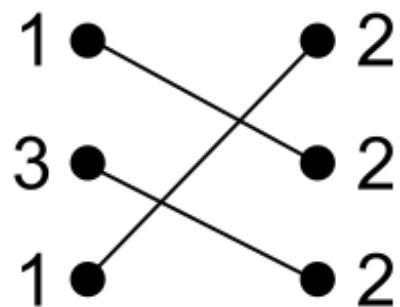


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

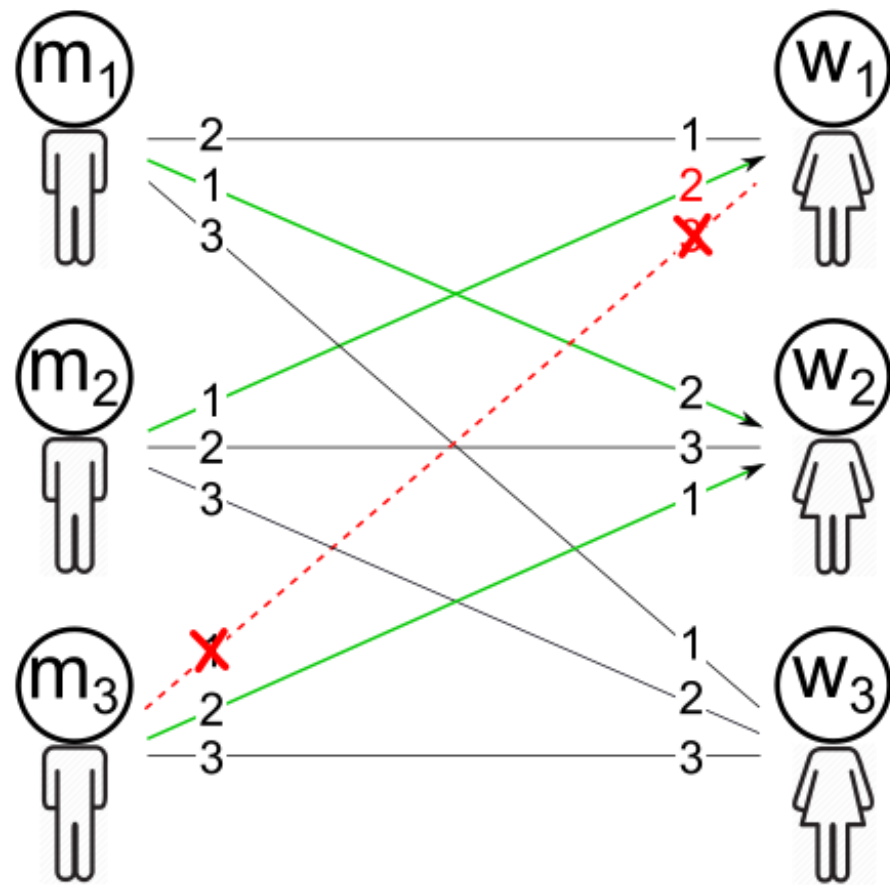
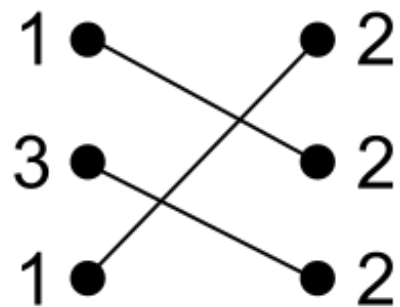


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

# DA algorithm is not strategyproof.

Outcome for true prefs



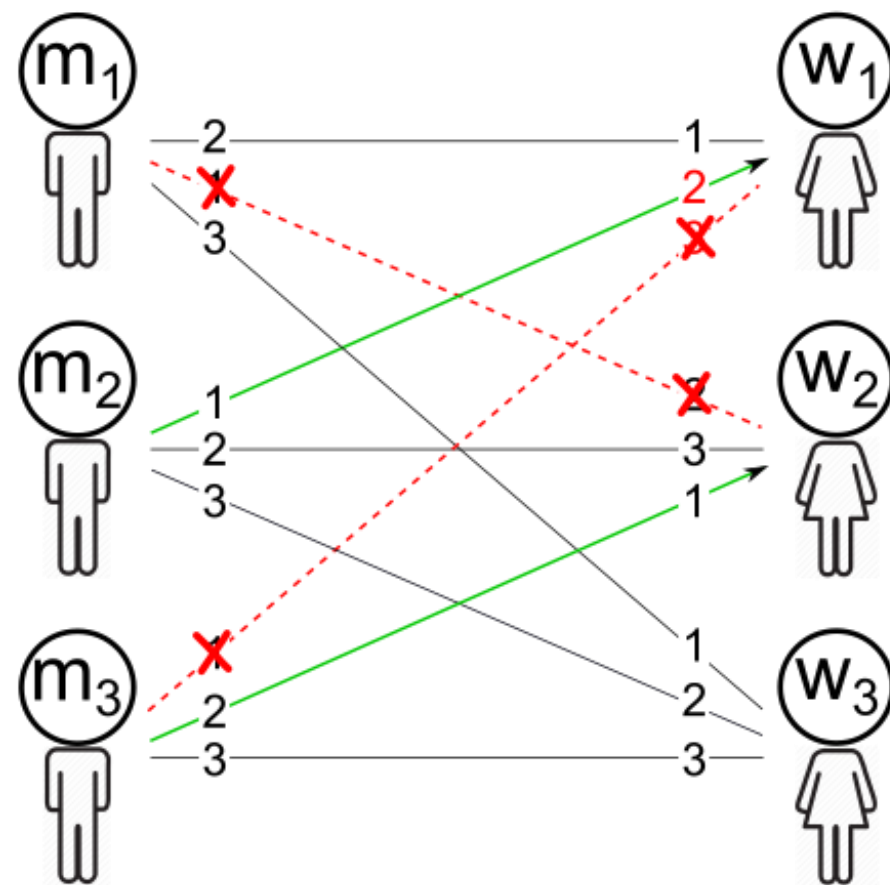
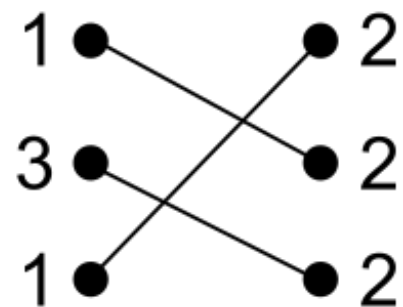
Can I get a better partner by misreporting my preferences?



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs



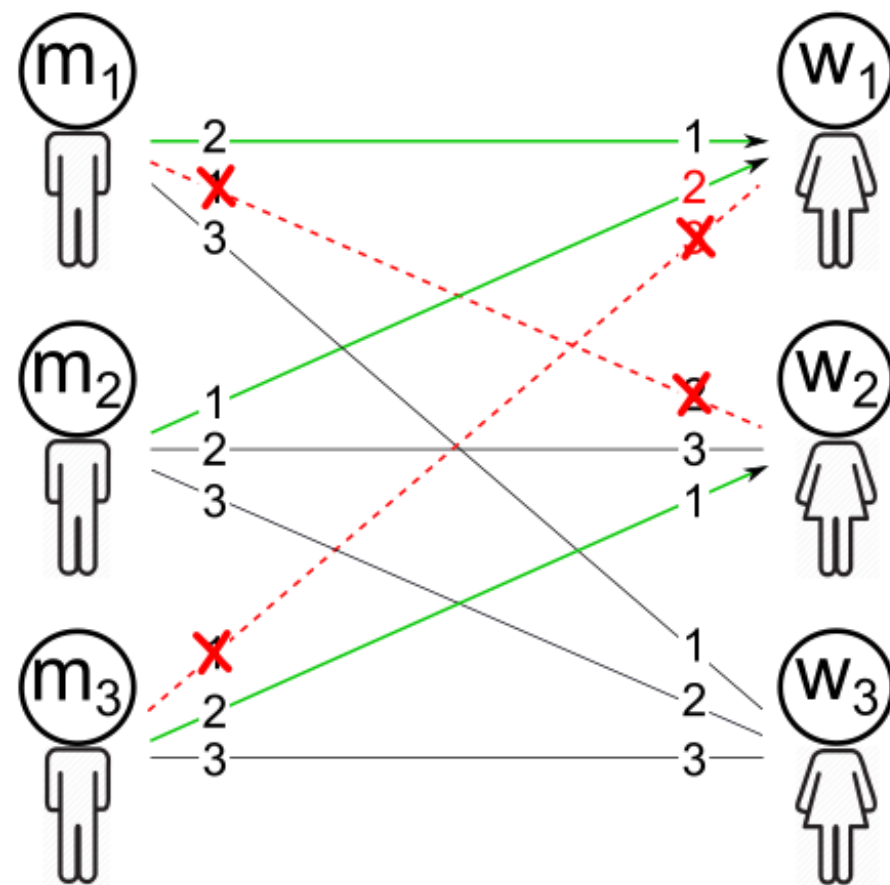
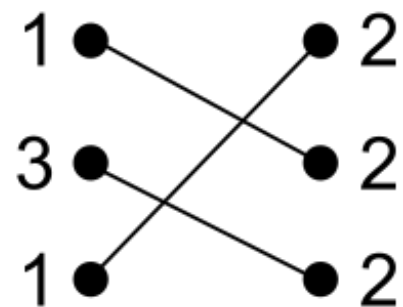
Can I get a better partner by misreporting my preferences?



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

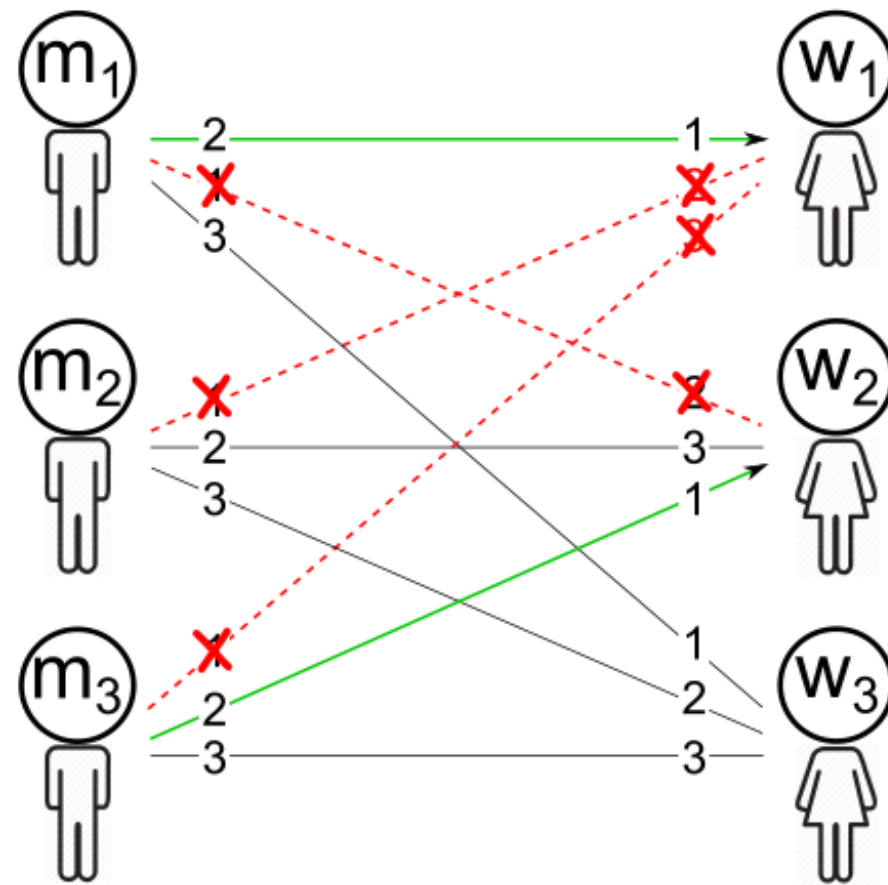
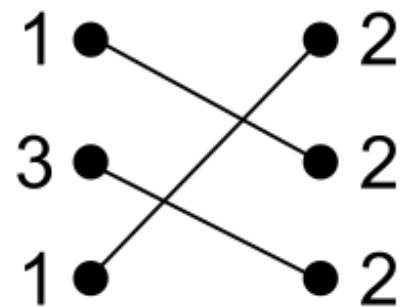


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

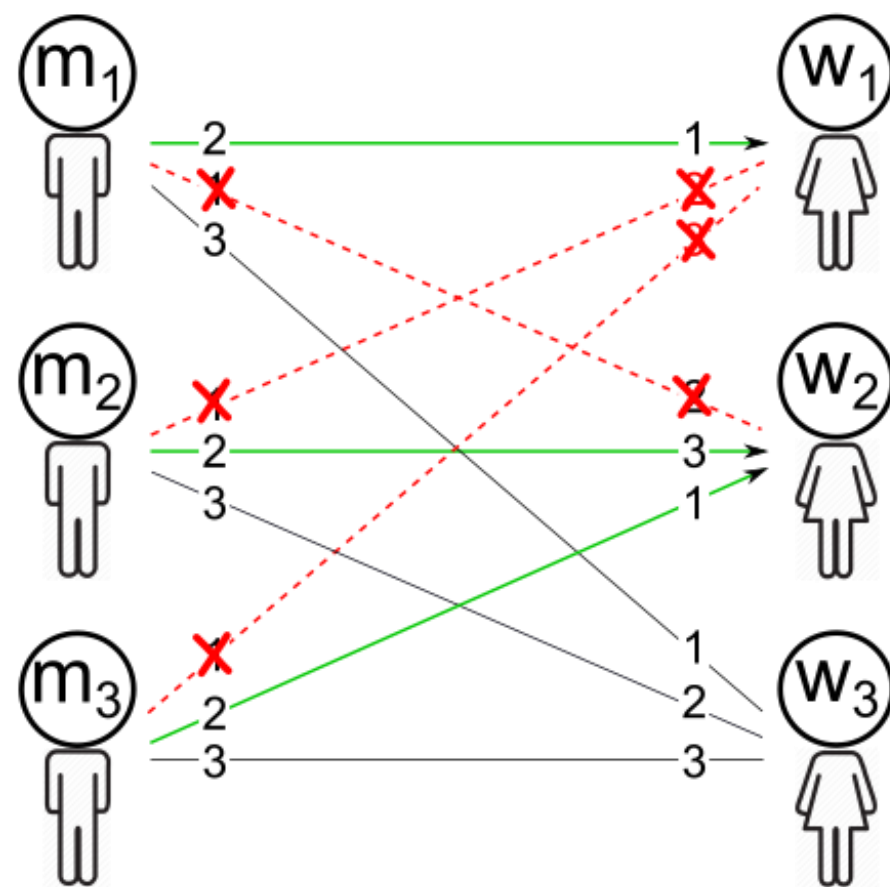
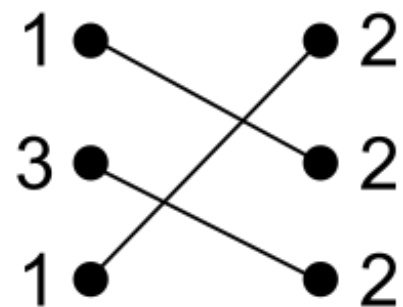


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

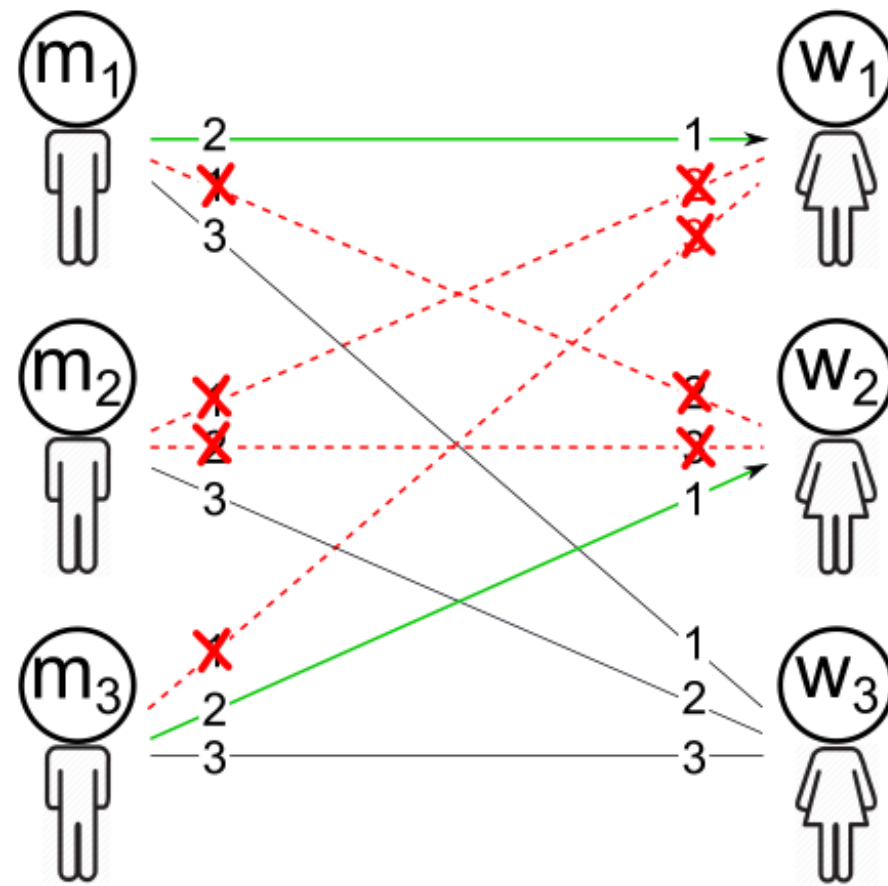
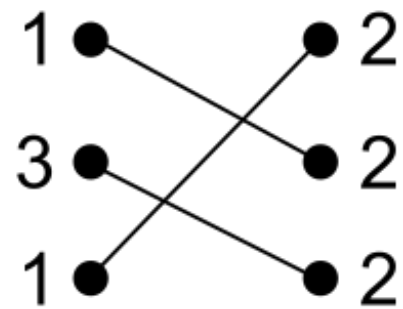


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

# DA algorithm is not strategyproof.

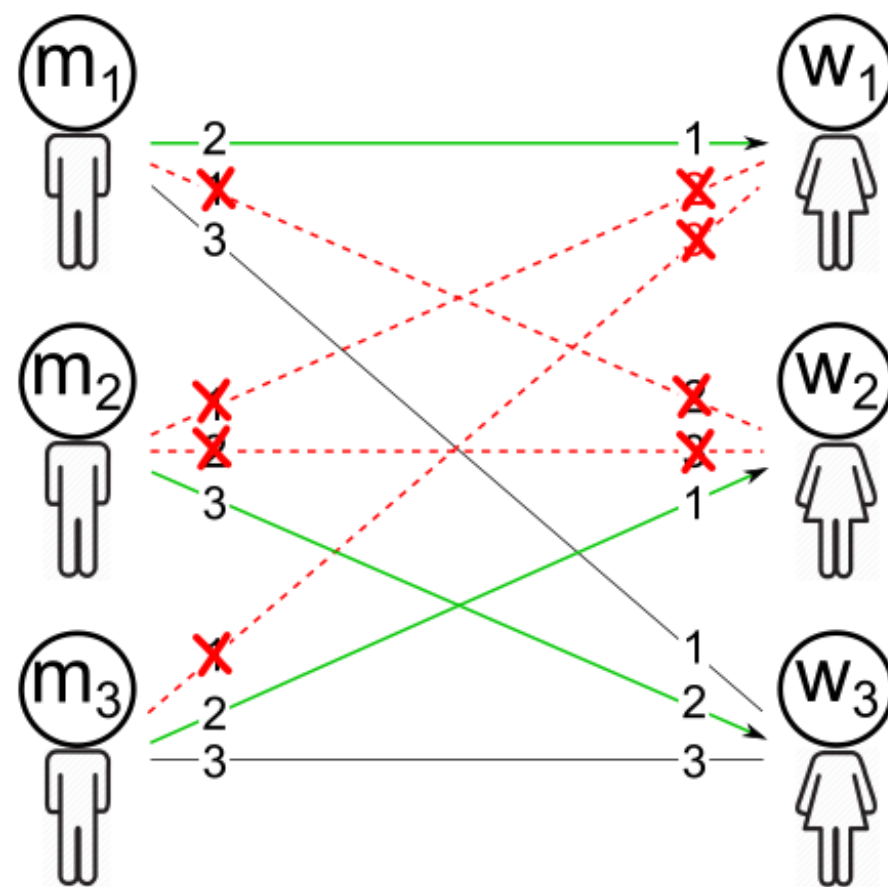
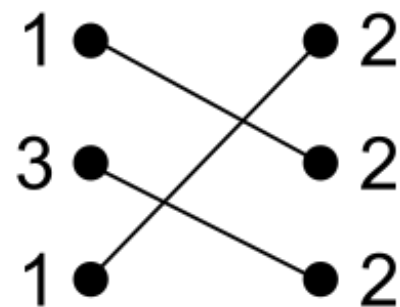
Outcome for true prefs



Can I get a better partner by misreporting my preferences?

DA algorithm is not strategyproof.

Outcome for true prefs

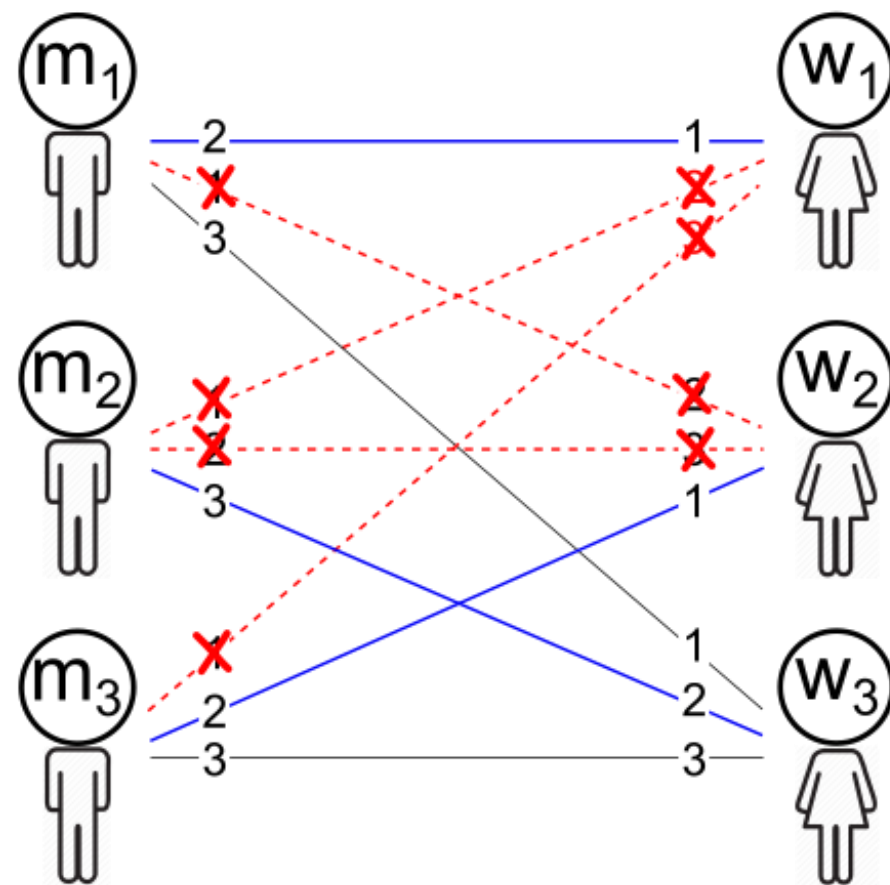
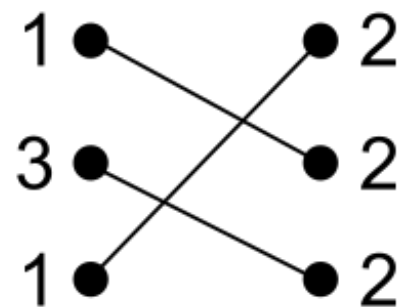


Can I get a better partner by misreporting my preferences?

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

## DA algorithm is not strategyproof.

Outcome for true prefs

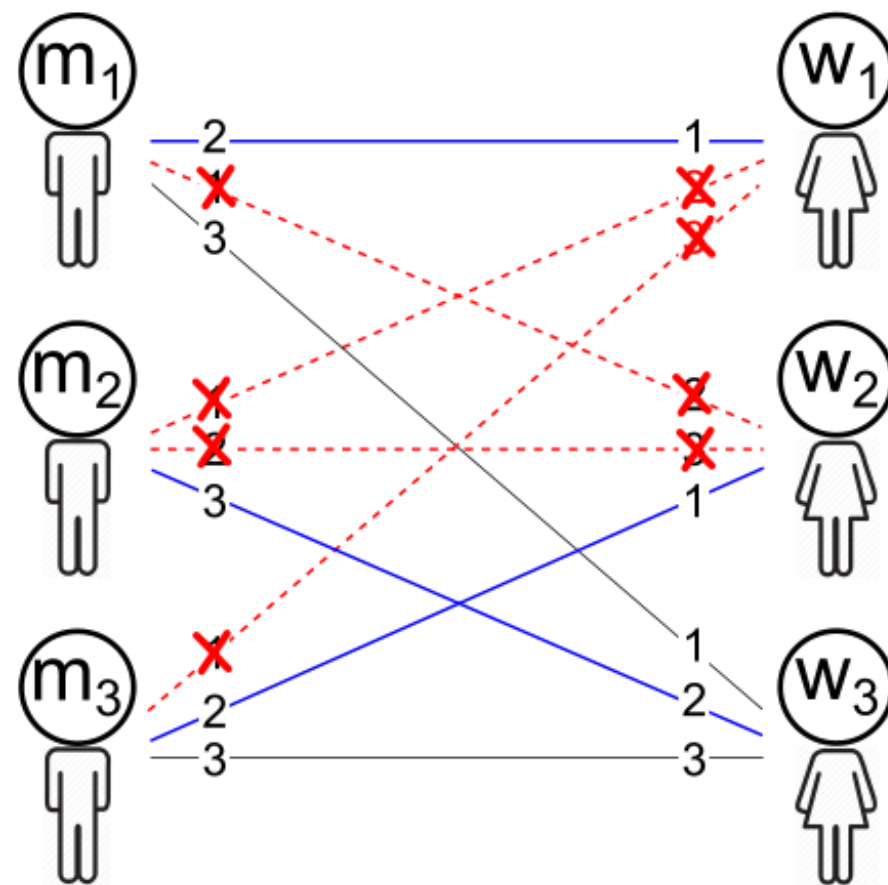
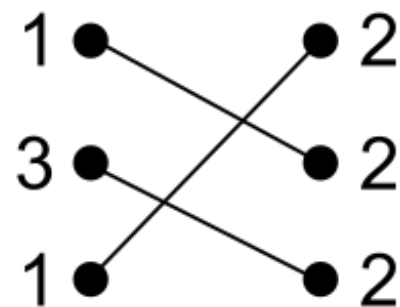


Can I get a better partner by misreporting my preferences?

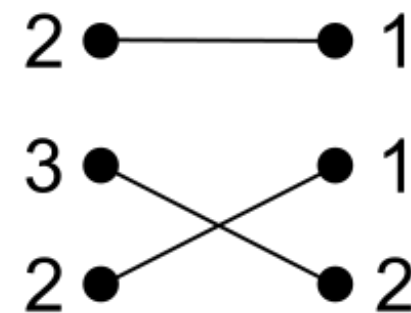
[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

# DA algorithm is not strategyproof.

Outcome for true prefs



Can I get a better partner by misreporting my preferences?

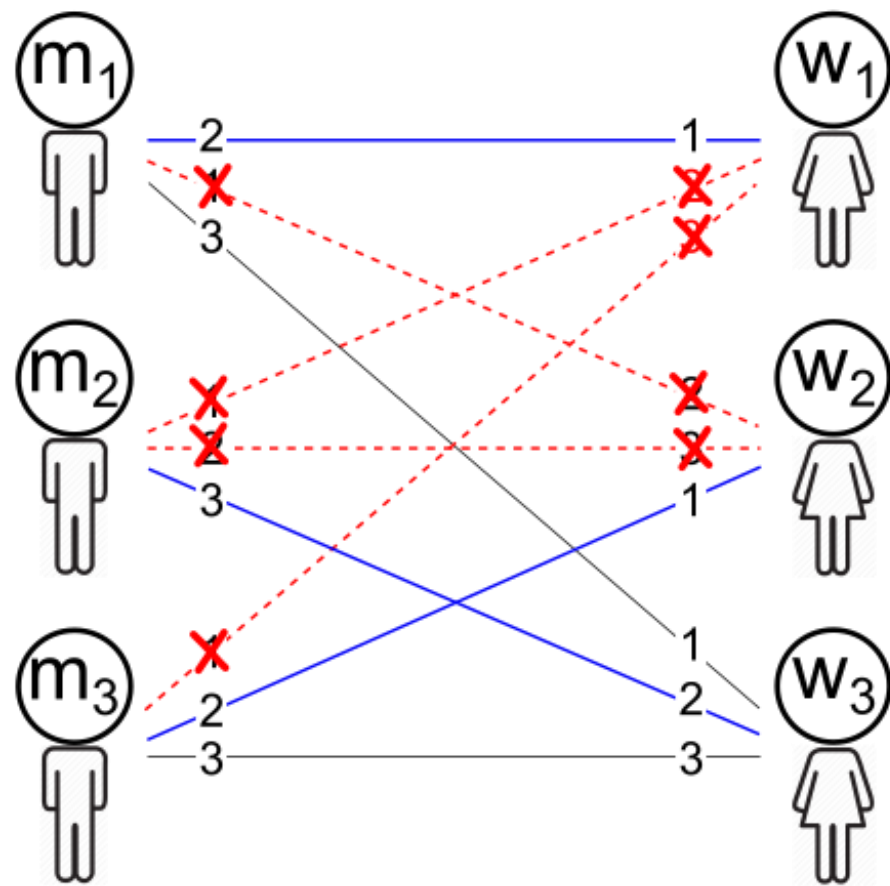
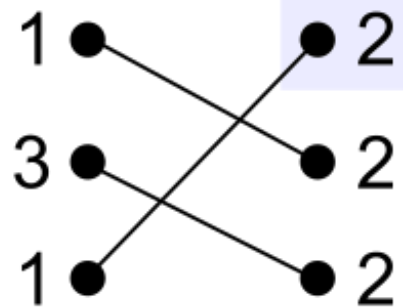




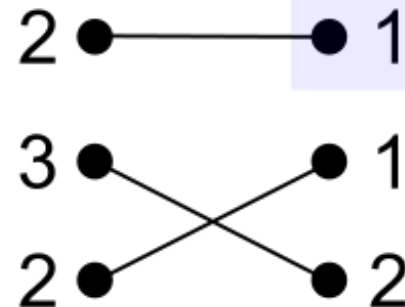
[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

# DA algorithm is not strategyproof.

Outcome for true prefs



Can I get a better partner by misreporting my preferences?

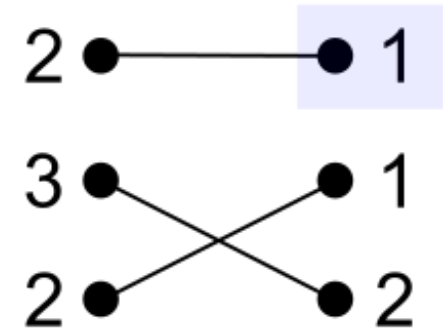
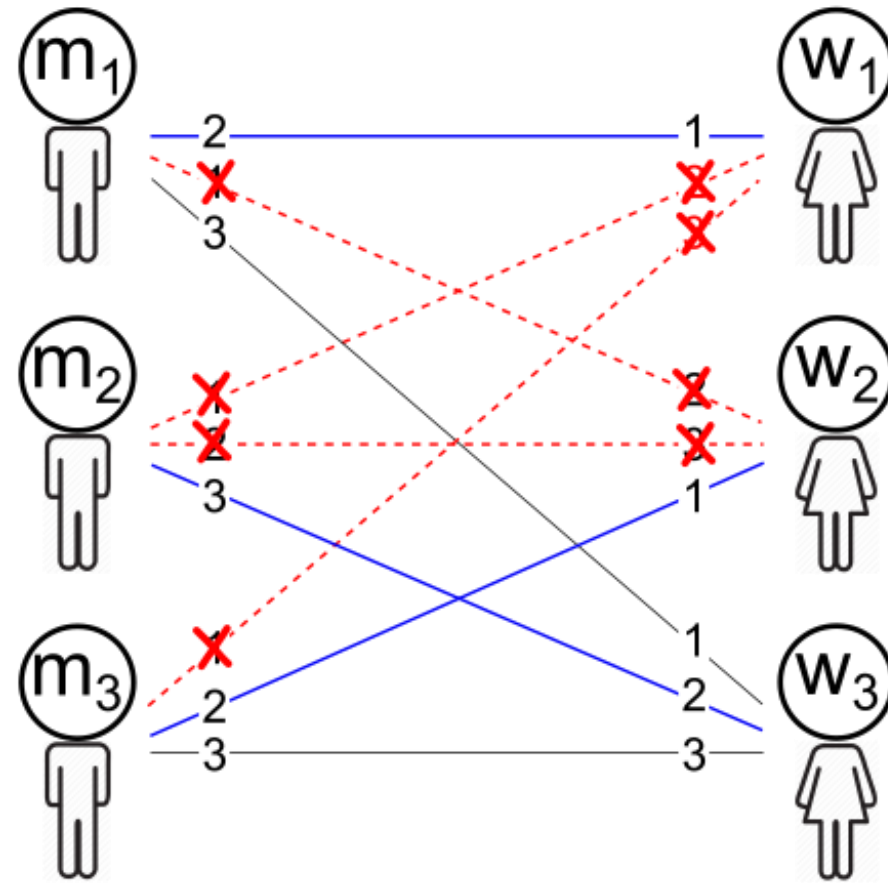
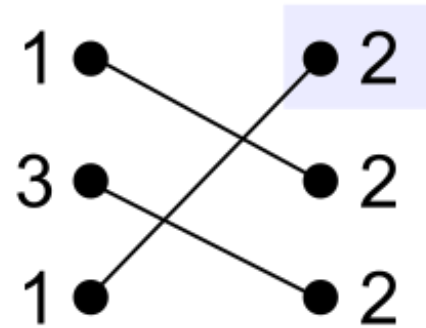




[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

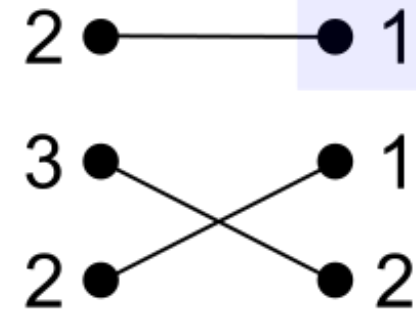
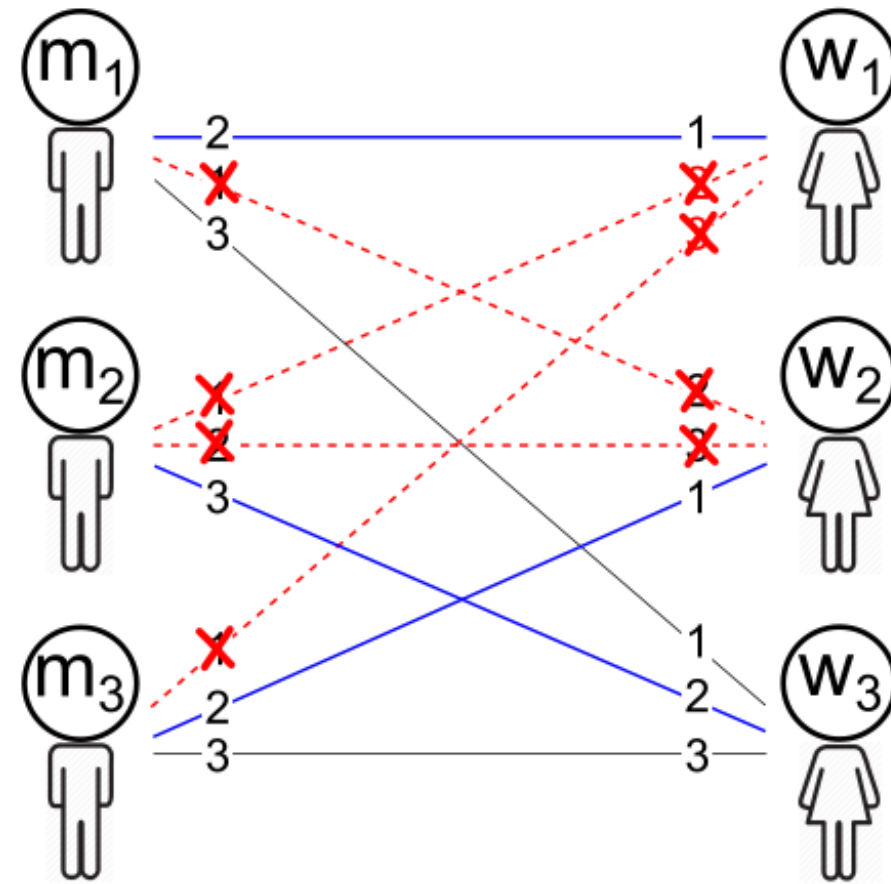
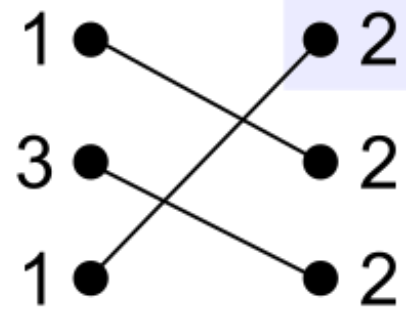
DA algorithm is not strategyproof.

Outcome for true prefs



DA algorithm is not strategyproof.

Outcome for true prefs



Any luck for the men?

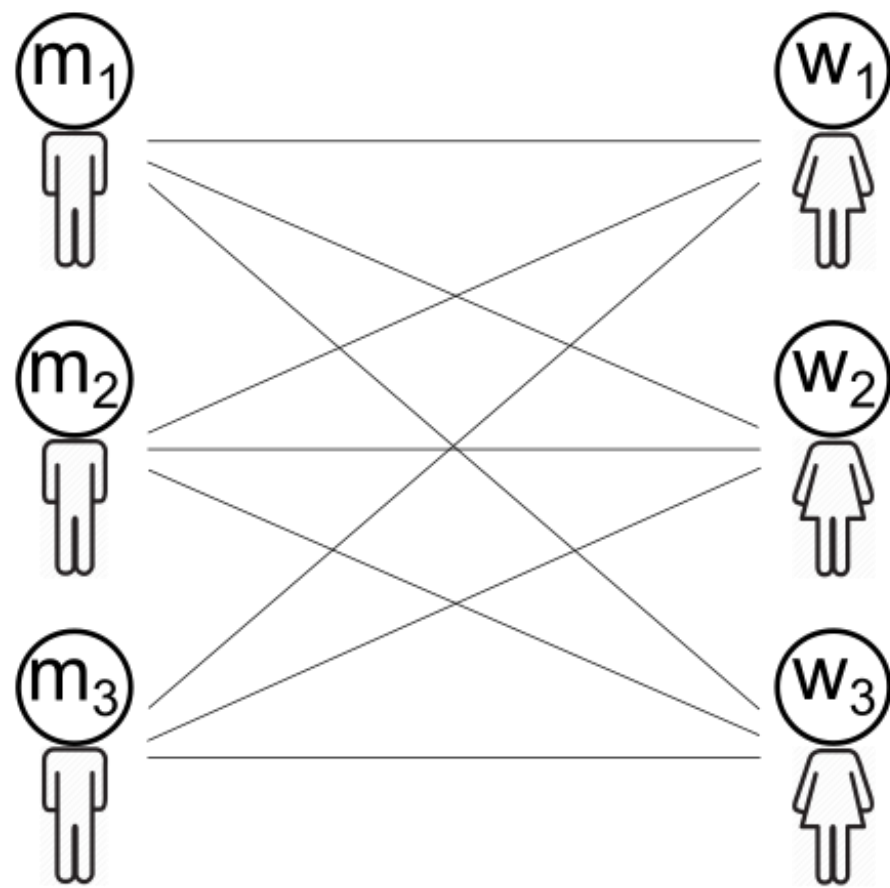
[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is strategyproof for the men.

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is strategyproof for the men.

Can I get a better partner by misreporting my preferences?



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

Can I get a better partner by misreporting my preferences?

Nope.

Nope.

Nope.

Nope.

Nope.

$m_1$

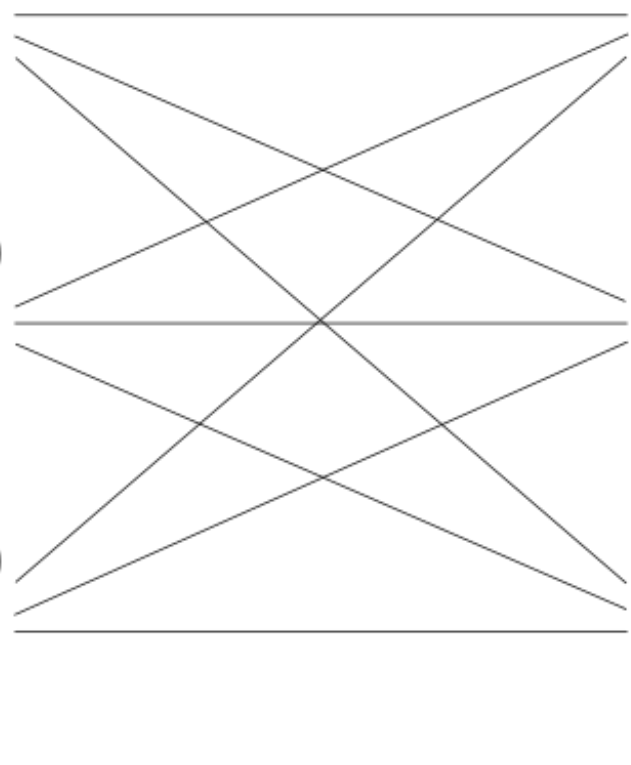
$m_2$

$m_3$

$w_1$

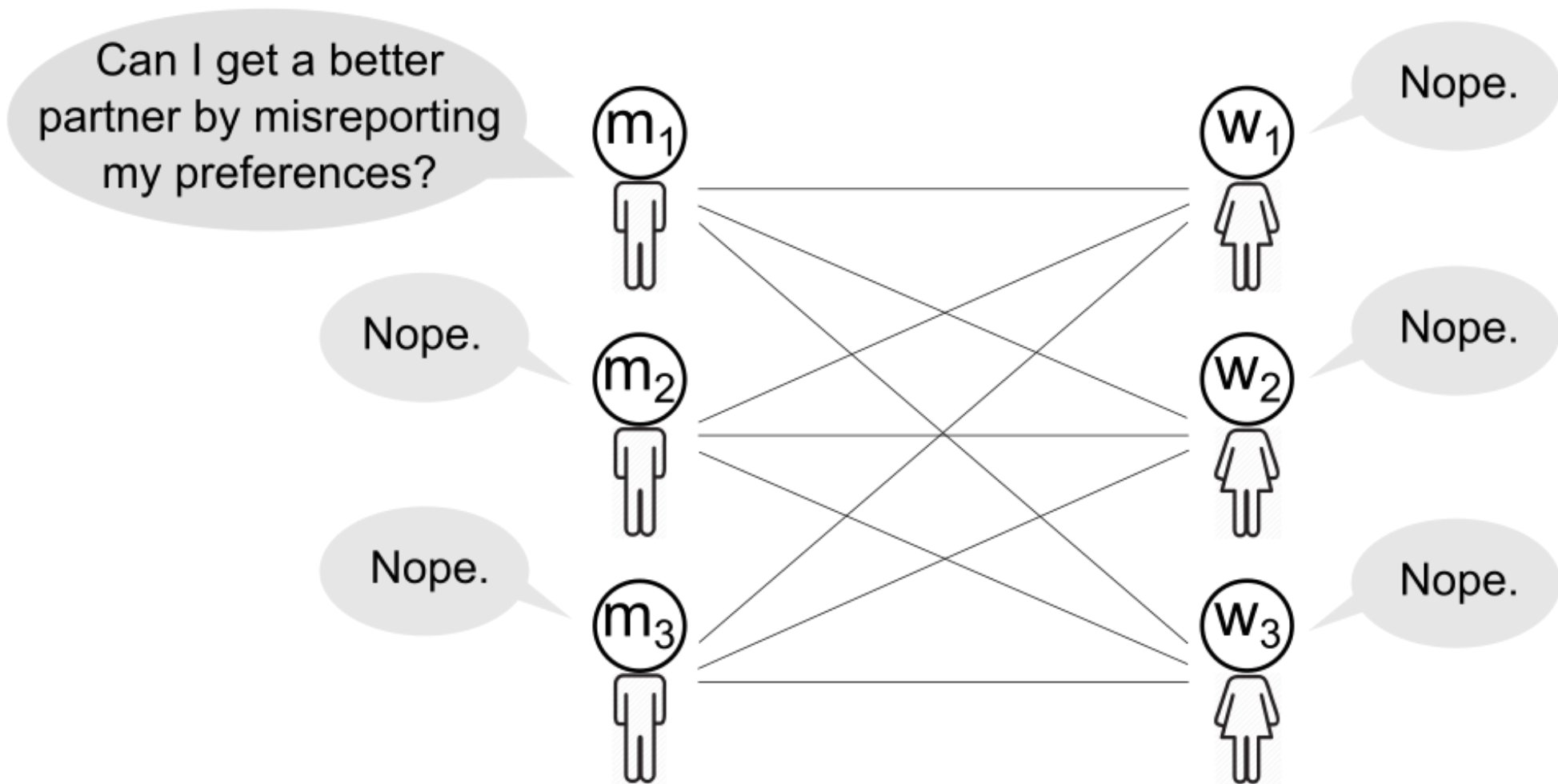
$w_2$

$w_3$



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**



Proof ~~later in today's lecture~~ in the lecture slides.

So, men can't cheat in the men-proposing DA algorithm but there **exists** an opportunity for a woman to manipulate.



So, men can't cheat in the men-proposing DA algorithm but there **exists** an opportunity for a woman to manipulate.

Is it possible to efficiently **compute** a beneficial misreport whenever one exists?

[Teo, Sethuraman and Tan, *Manag. Sci.* 2001; Vaish and Garg, *IJCAI* 2017]

An optimal manipulation for a woman can be computed  
in polynomial time.

[Teo, Sethuraman and Tan, *Manag. Sci.* 2001; Vaish and Garg, *IJCAI* 2017]

An optimal manipulation for a woman can be computed  
in polynomial time.

We will use a structural result about optimal manipulation strategies.

Any optimal manipulation for a woman can also be achieved by an "inconspicuous" misreport.

Any optimal manipulation for a woman can also be achieved by an "inconspicuous" misreport.

True list of woman  $w$ :  $m_1 > m_2 > m_3 > m_4 > \boxed{m_5} > m_6 > m_7 > m_8$

Any optimal manipulation for a woman can also be achieved by an "inconspicuous" misreport.

True list of woman  $w$ :  $m_1 > m_2 > m_3 > m_4 > \boxed{m_5} > m_6 > m_7 > m_8$

An optimal manipulation:  $\boxed{m_2} > m_4 > m_1 > m_8 > m_6 > m_3 > m_5 > m_7$

Any optimal manipulation for a woman can also be achieved by an "inconspicuous" misreport.

True list of woman  $w$ :  $m_1 > m_2 > m_3 > m_4 > \boxed{m_5} > m_6 > m_7 > m_8$

An optimal manipulation:  $\boxed{m_2} > m_4 > m_1 > m_8 > m_6 > m_3 > m_5 > m_7$

An *inconspicuous* misreport that is also optimal for  $w$ :  $m_1 > \boxed{m_2} > m_6 > m_3 > m_4 > m_5 > m_7 > m_8$

Any optimal manipulation for a woman can also be achieved by an "inconspicuous" misreport.

True list of woman  $w$ :  $m_1 > m_2 > m_3 > m_4 > \boxed{m_5} > m_6 > m_7 > m_8$

An optimal manipulation:  $\boxed{m_2} > m_4 > m_1 > m_8 > m_6 > m_3 > m_5 > m_7$

An *inconspicuous* misreport that is also optimal for  $w$ :  $m_1 > \boxed{m_2} > m_6 > m_3 > m_4 > m_5 > m_7 > m_8$



Any optimal manipulation for a woman can also be achieved by an "inconspicuous" misreport.

True list of woman  $w$ :  $m_1 > m_2 > m_3 > m_4 > \boxed{m_5} > m_6 > m_7 > m_8$

An optimal manipulation:  $\boxed{m_2} > m_4 > m_1 > m_8 > m_6 > m_3 > m_5 > m_7$

An *inconspicuous* misreport that is also optimal for  $w$ :  $m_1 > \boxed{m_2} > m_6 > m_3 > m_4 > m_5 > m_7 > m_8$

[Teo, Sethuraman and Tan, *Manag. Sci.* 2001; Vaish and Garg, *IJCAI* 2017]

An optimal manipulation for a woman can be computed in polynomial time.

But what about stability (w.r.t. true preferences)?

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

We will use the following observation:

Suppose a woman promotes a man  $m$  in her list and no other changes are made.

If  $m$  proposed to her during DA on the old profile, then he proposes to her during DA on the new profile.

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

We will use the following observation:

Suppose a woman promotes a man  $m$  in her list and no other changes are made.

If  $m$  proposed to her during DA on the old profile, then he proposes to her during DA on the new profile.

*Idea:* Any deviation between old and new runs of the DA must involve tentative acceptance/rejection of  $m$ , but that can happen only after  $m$  proposes.

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$$P = (P_{-w}, P_w)$$

$P'$  = profile with **manipulated** preferences

$$P' = (P_{-w}, P'_w)$$



The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$$P = (P_{-w}, P_w)$$

$$X = DA(P)$$

$P'$  = profile with **manipulated** preferences

$$P' = (P_{-w}, P'_w)$$

$$X' = DA(P')$$

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$$P = (P_{-w}, P_w)$$

$$X = DA(P)$$

$P'$  = profile with **manipulated** preferences

$$P' = (P_{-w}, P'_w)$$

$$X' = DA(P')$$

Suppose, for contradiction, that  $X'$  is not stable w.r.t.  $P$ .

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$$P = (P_{-w}, P_w)$$

$$X = DA(P)$$

$P'$  = profile with **manipulated** preferences

$$P' = (P_{-w}, P'_w)$$

$$X' = DA(P')$$

Suppose, for contradiction, that  $X'$  is not stable w.r.t.  $P$ .

Then, there must be a pair  $(m, w')$  that blocks  $X'$  w.r.t.  $P$ .

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$$P = (P_{-w}, P_w)$$

$$X = DA(P)$$

$P'$  = profile with **manipulated** preferences

$$P' = (P_{-w}, P'_w)$$

$$X' = DA(P')$$

Suppose, for contradiction, that  $X'$  is not stable w.r.t.  $P$ .

Then, there must be a pair  $(m, w')$  that blocks  $X'$  w.r.t.  $P$ .

It must be that  $w' = w$ .

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$$P = (P_{-w}, P_w)$$

$$X = DA(P)$$

$P'$  = profile with **manipulated** preferences

$$P' = (P_{-w}, P'_w)$$

$$X' = DA(P')$$

Suppose, for contradiction, that  $X'$  is not stable w.r.t.  $P$ .

Then, there must be a pair  $(m, w')$  that blocks  $X'$  w.r.t.  $P$ .

It must be that  $w' = w$ .

If  $w' \neq w$ , then  $m$  and  $w'$  are both truthful and will block  $X'$  w.r.t.  $P'$ ---contradicting the stability of DA.

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$$P = (P_{-w}, P_w)$$

$$X = DA(P)$$

$P'$  = profile with **manipulated** preferences

$$P' = (P_{-w}, P'_w)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .

$P_m$	$P_w$
•	•
•	•
w	m
•	•
•	•
$X'(m)$	$X'(w)$
•	•
•	•

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .

$P_m$	$P_w$	$P'_m$	$P'_w$
•	•	•	•
•	•	•	•
w	m	w	$X'(w)$
•	•	•	•
•	•	•	•
$X'(m)$	$X'(w)$	$X'(m)$	m
•	•	•	•
•	•	•	•



The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .

$P_m$	$P_w$	$P'_m$	$P'_w$
•	•	•	•
•	•	•	•
w	m	w	$X'(w)$
•	•	•	•
•	•	•	•
$X'(m)$	$X'(w)$	$X'(m)$	m
•	•	•	•
•	•	•	•

$m$  must propose to  $w$  during  $DA(P')$

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

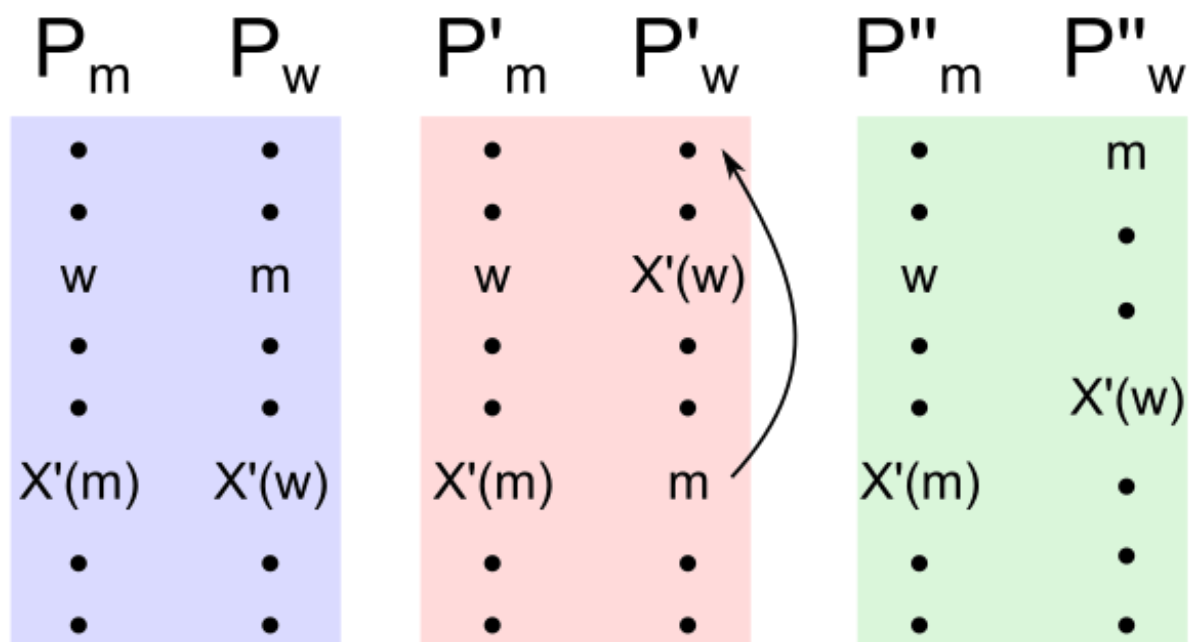
$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .



$m$  must propose to  $w$  during  $DA(P')$

## Recall:

Suppose a woman promotes a man  $m$  in her list  
and no other changes are made.

If  $m$  proposed to her during DA on the old profile,  
then he proposes to her during DA on the new profile.

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

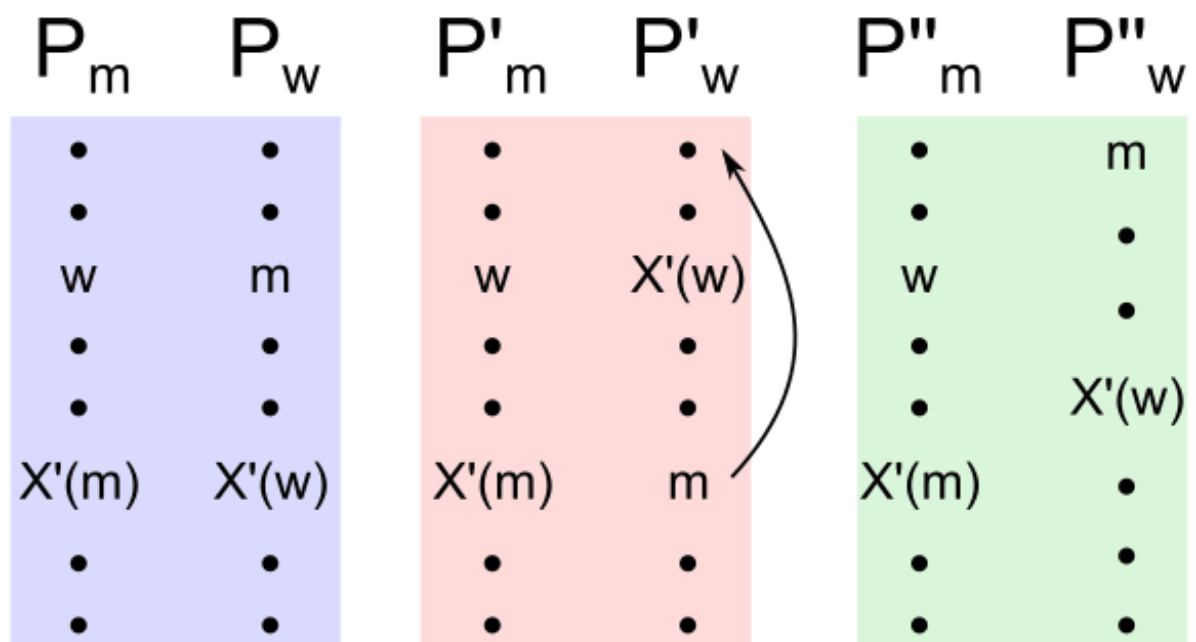
$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .



$m$  must propose to  $w$  during  $DA(P')$

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

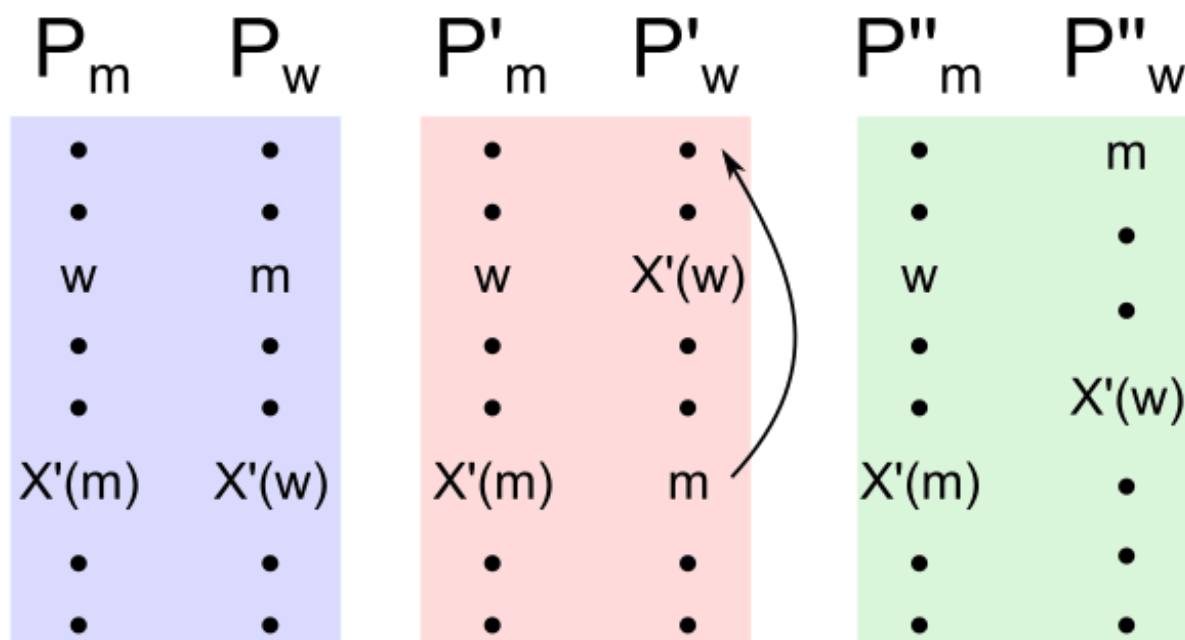
$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .



$m$  must propose to  $w$  during  $DA(P')$   
 $\Rightarrow m$  must propose to  $w$  during  $DA(P'')$

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

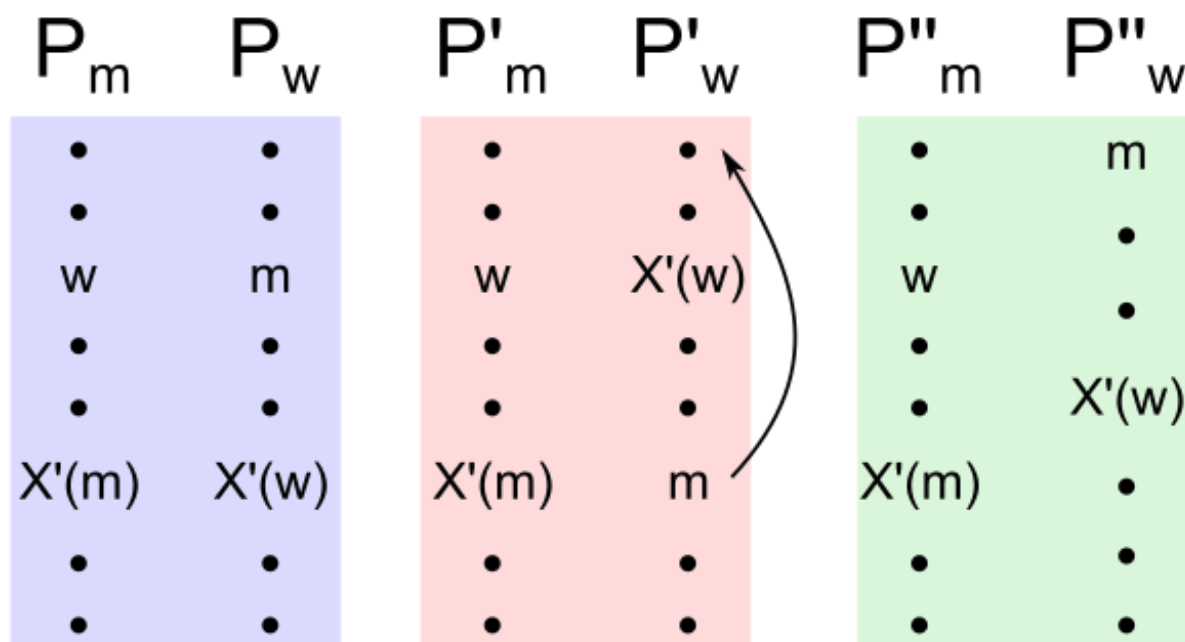
$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .



$m$  must propose to  $w$  during  $DA(P')$   
 $\Rightarrow m$  must propose to  $w$  during  $DA(P'')$   
 $\Rightarrow X''(w) = m$ , where  $X'' = DA(P'')$

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

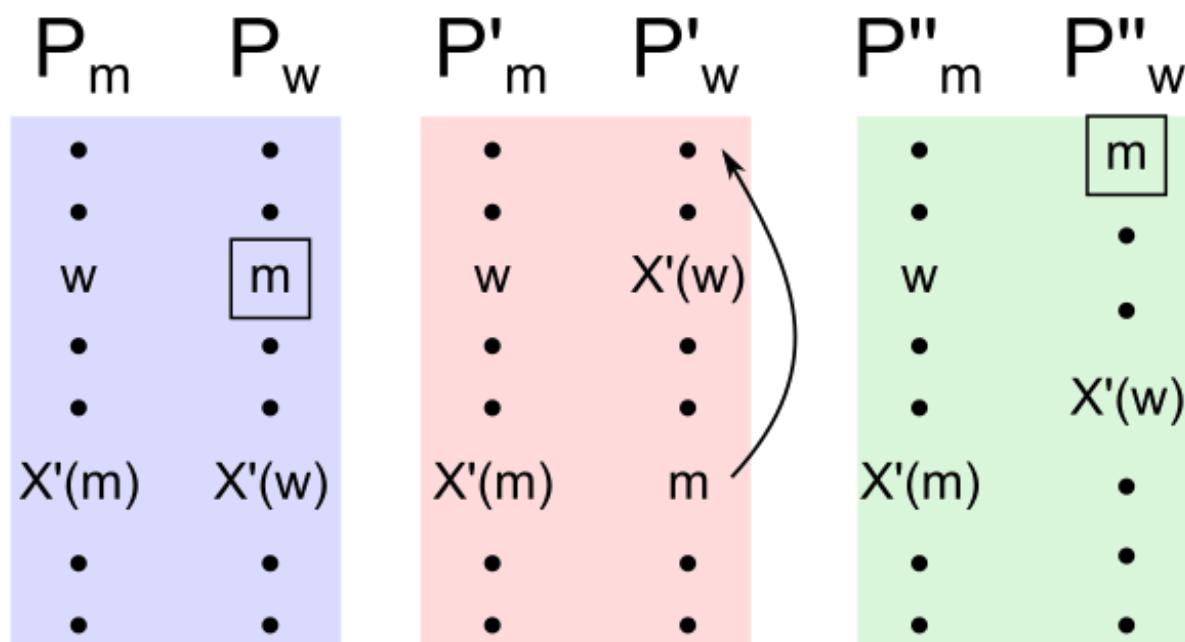
$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .



$m$  must propose to  $w$  during  $DA(P')$   
 $\Rightarrow m$  must propose to  $w$  during  $DA(P'')$   
 $\Rightarrow X''(w) = m$ , where  $X'' = DA(P'')$



The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

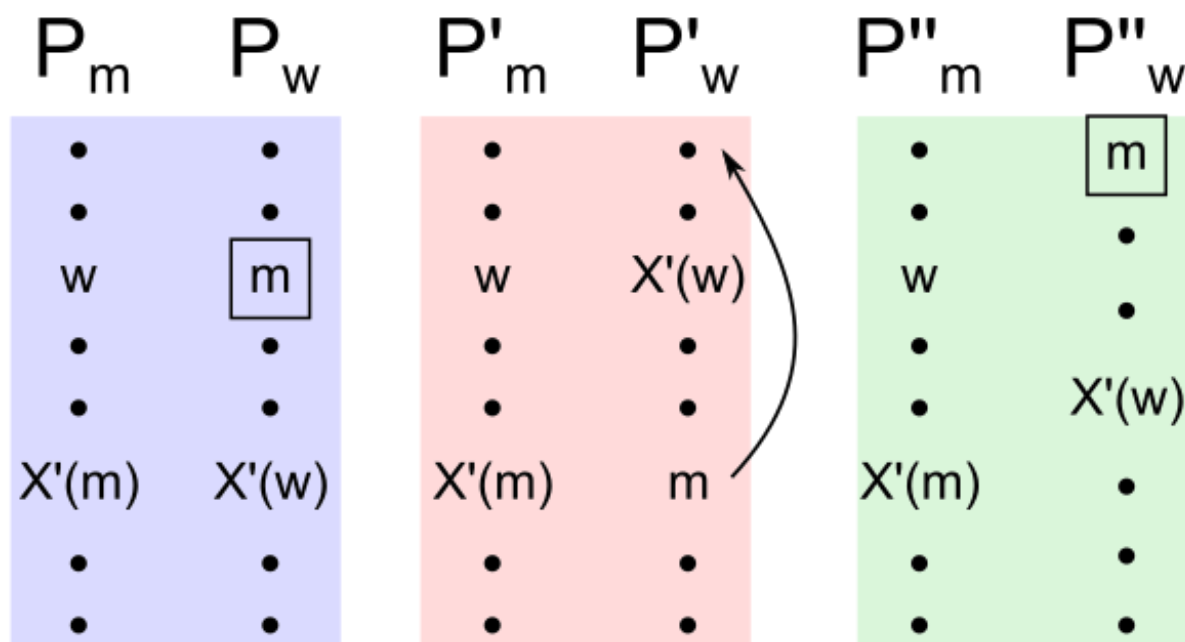
$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .



$m$  must propose to  $w$  during  $DA(P')$   
 $\Rightarrow m$  must propose to  $w$  during  $DA(P'')$   
 $\Rightarrow X''(w) = m$ , where  $X'' = DA(P'')$   
 But then,  $P''_w$  gives  $w$  a better partner than under her optimal strategy!



The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

$P$  = profile with **true** preferences

$P'$  = profile with **manipulated** preferences

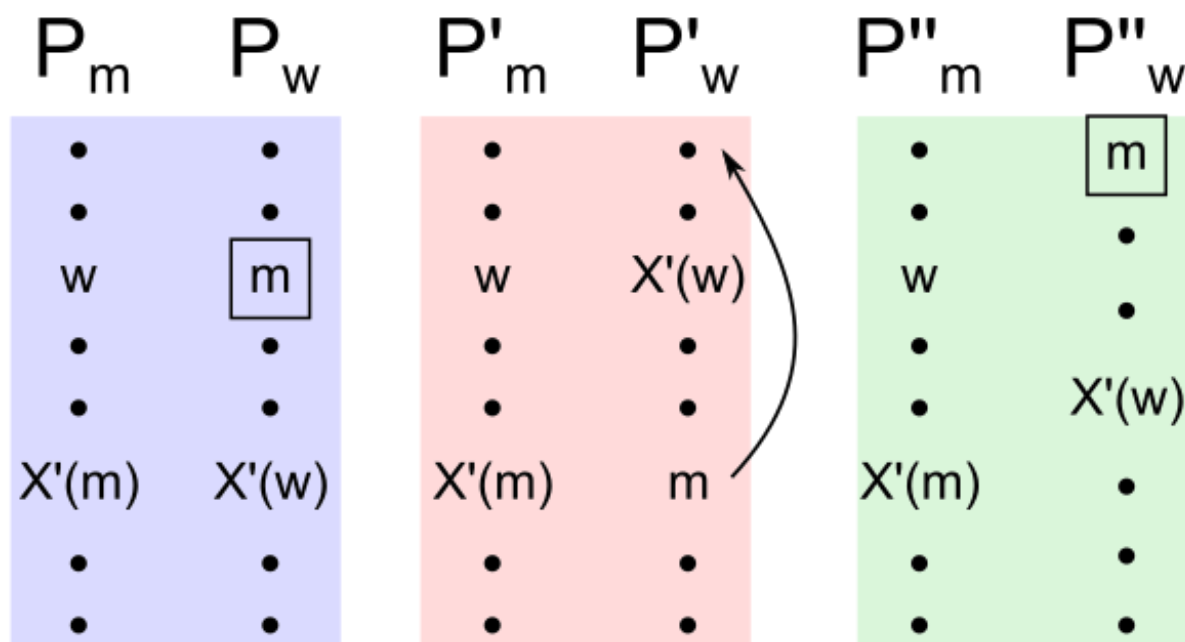
$$P = (P_{-w}, P_w)$$

$$P' = (P_{-w}, P'_w)$$

$$X = DA(P)$$

$$X' = DA(P')$$

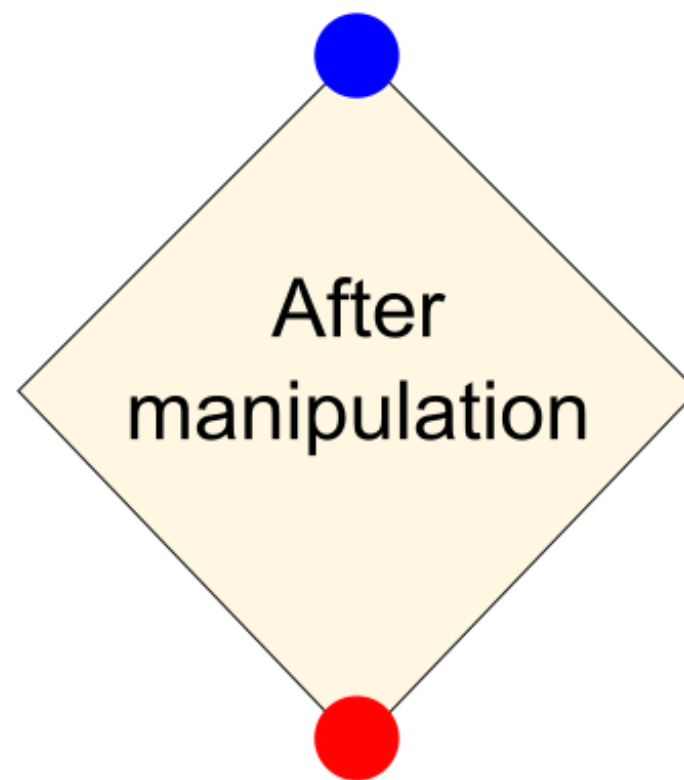
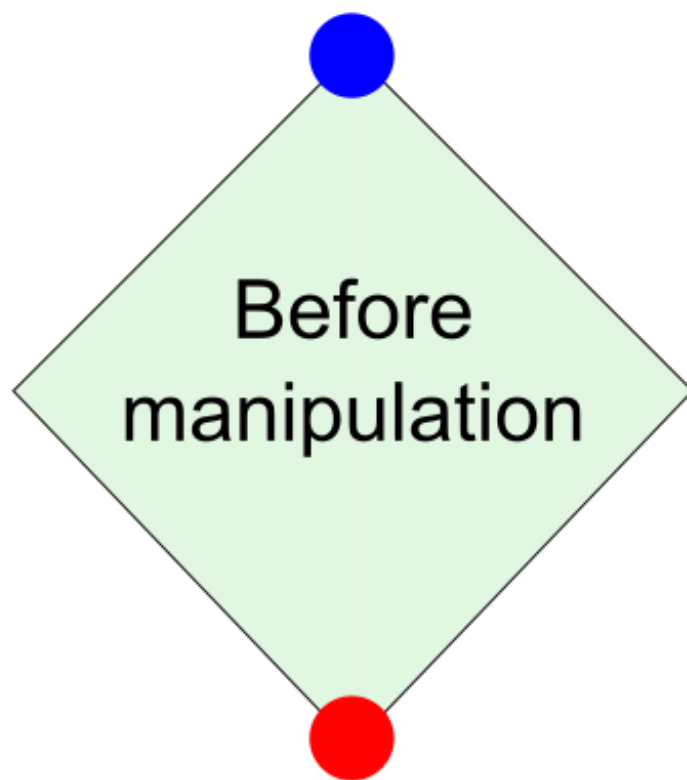
So,  $(m, w)$  blocks  $X'$  w.r.t.  $P$ .



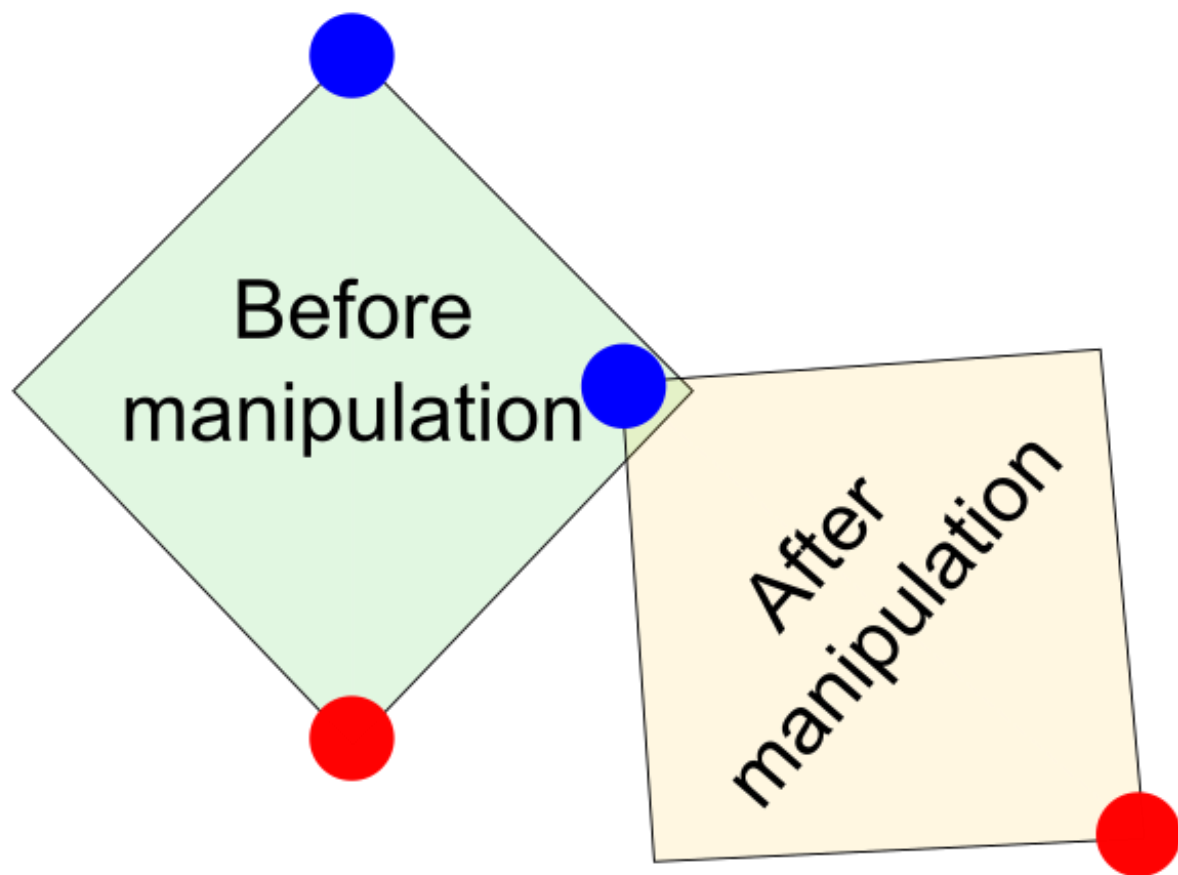
$m$  must propose to  $w$  during  $DA(P')$   
 $\Rightarrow m$  must propose to  $w$  during  $DA(P'')$   
 $\Rightarrow X''(w) = m$ , where  $X'' = DA(P'')$   
 But then,  $P''_w$  gives  $w$  a better partner than under her optimal strategy!

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.



The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.



The DA matching after optimal manipulation by a woman is stable with respect to the true preferences.

*Stable marriages are manipulable,  
but optimally manipulated marriages are stable.*

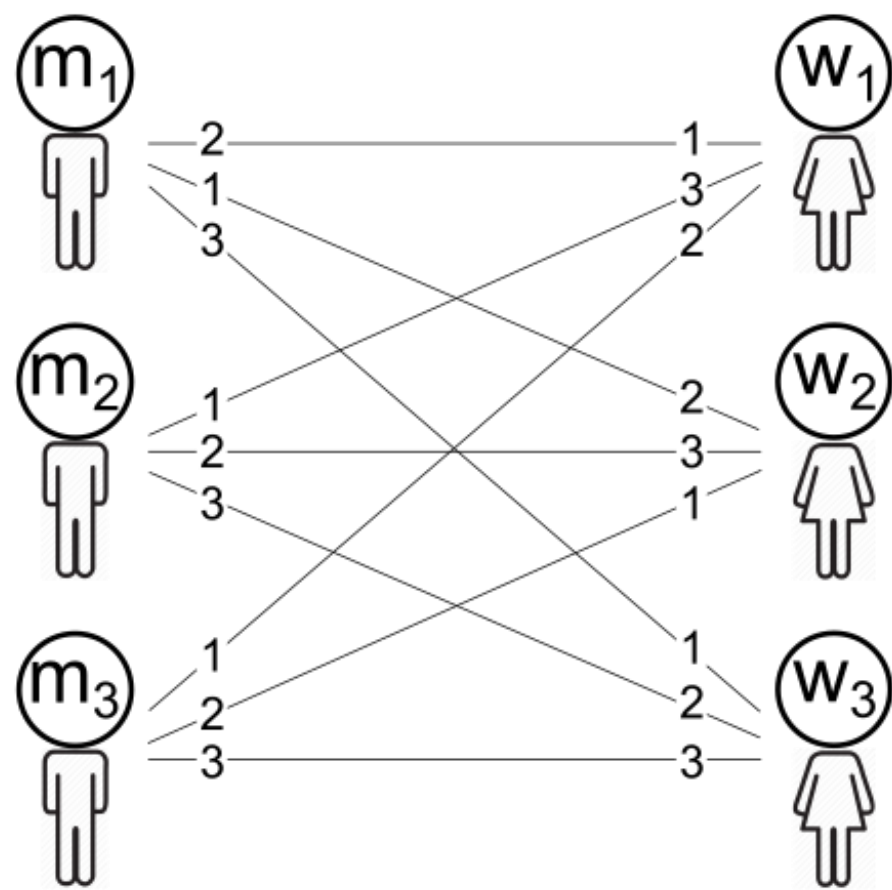
DA fails strategyproofness---too bad!

Let's think of a different stable matching algorithm that is truthful.

[Roth, *MOR* 1982]

No stable matching procedure can be strategyproof.

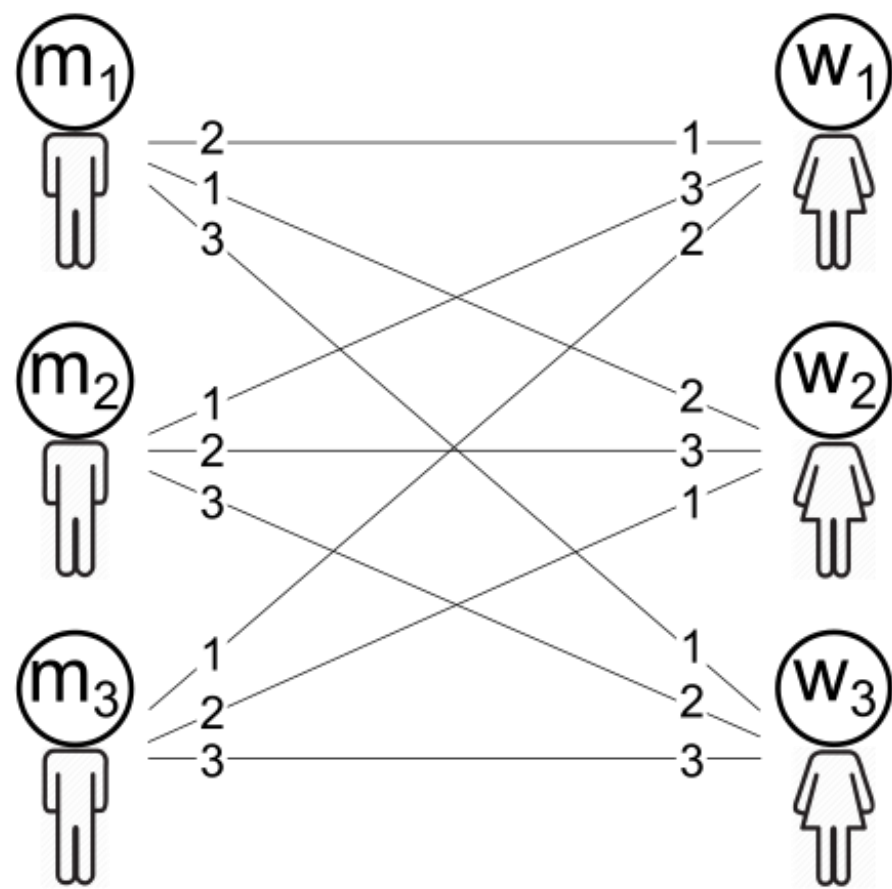
No stable matching procedure can be strategyproof.



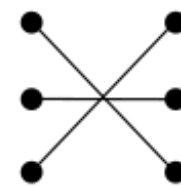
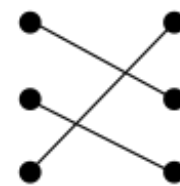
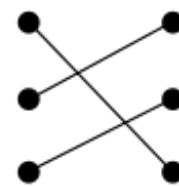
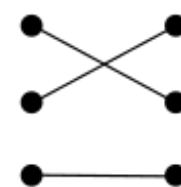
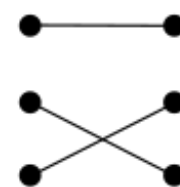
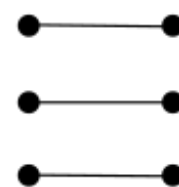
Instance  $I_0$



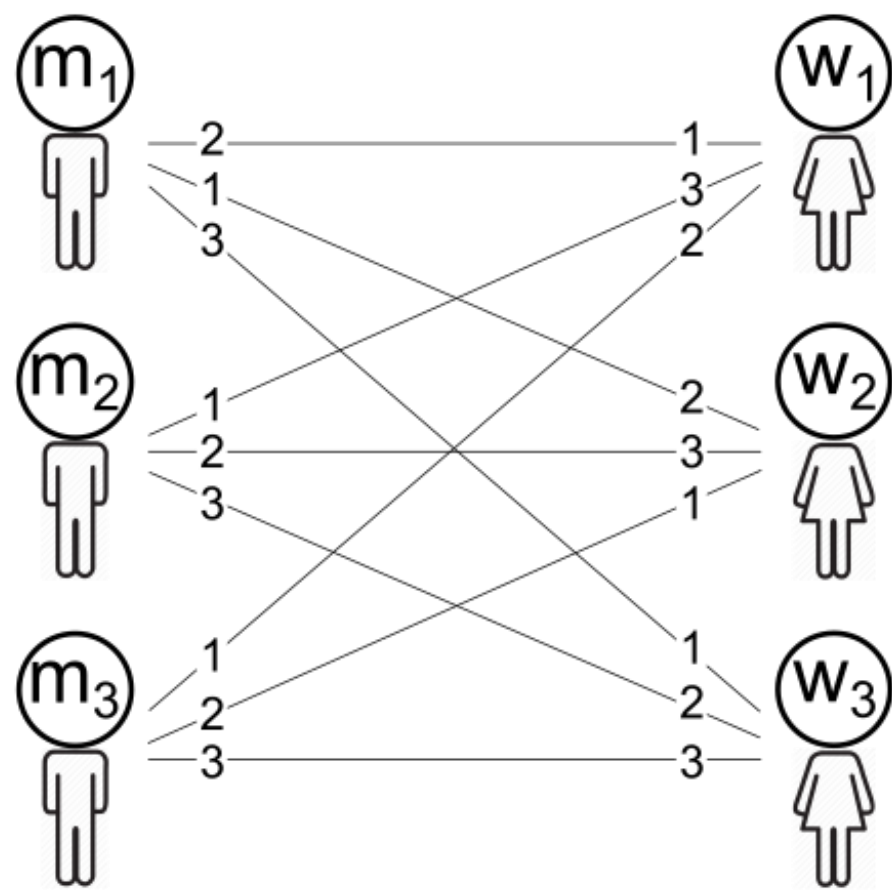
No stable matching procedure can be strategyproof.



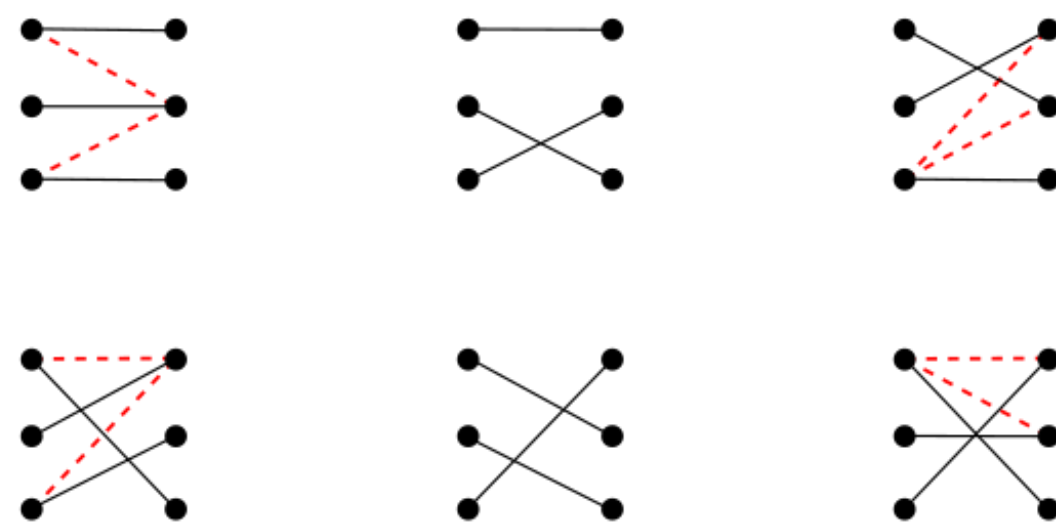
Instance  $I_0$



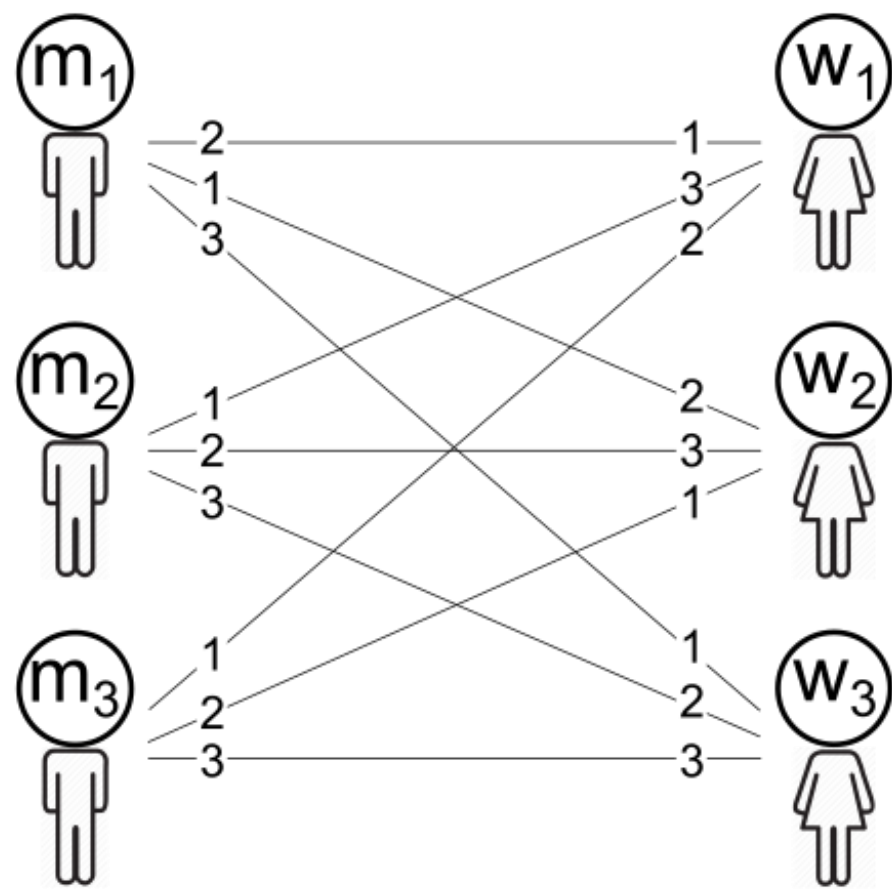
No stable matching procedure can be strategyproof.



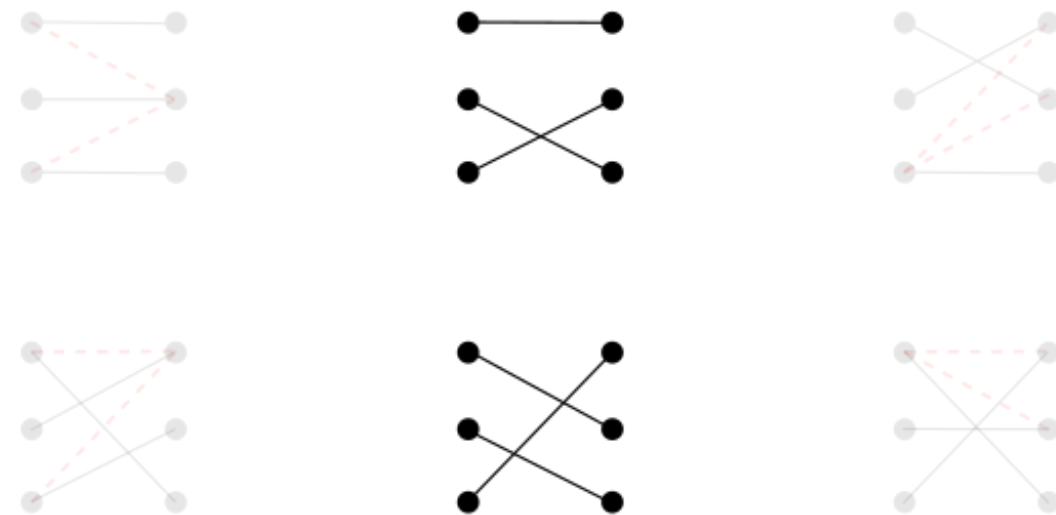
Instance  $I_0$



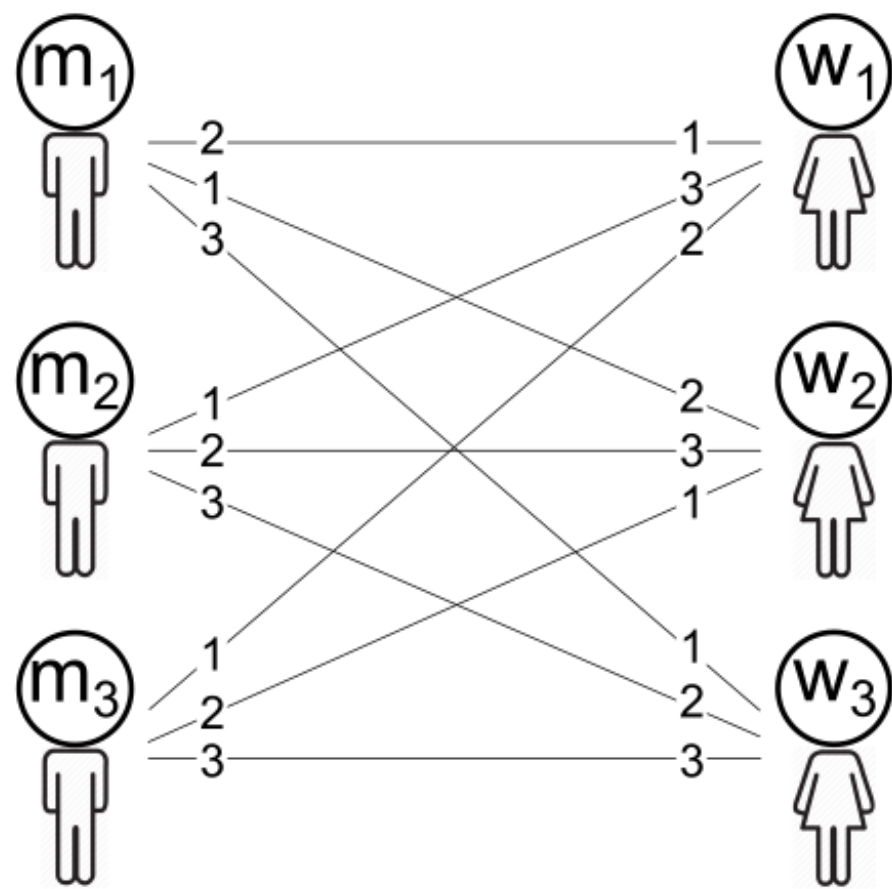
No stable matching procedure can be strategyproof.



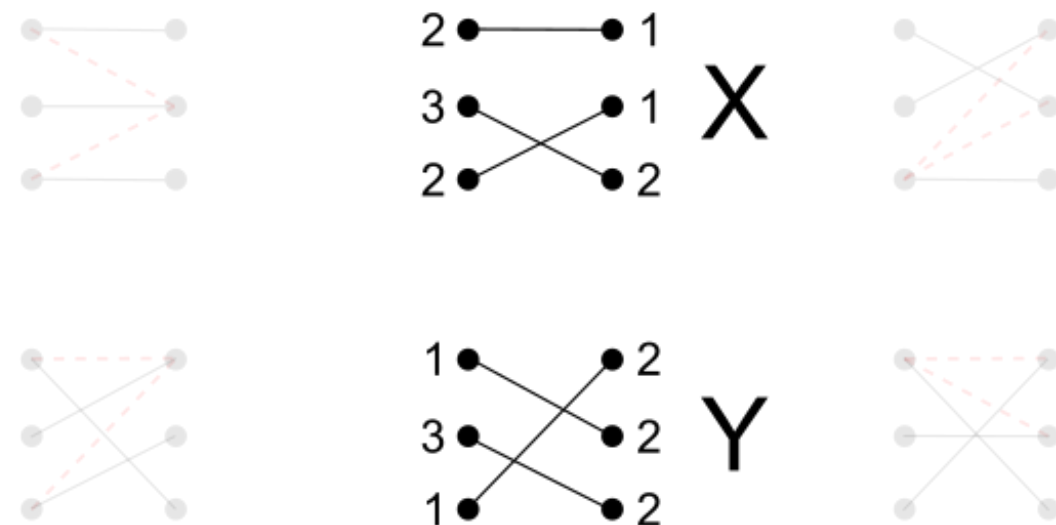
Instance  $I_0$



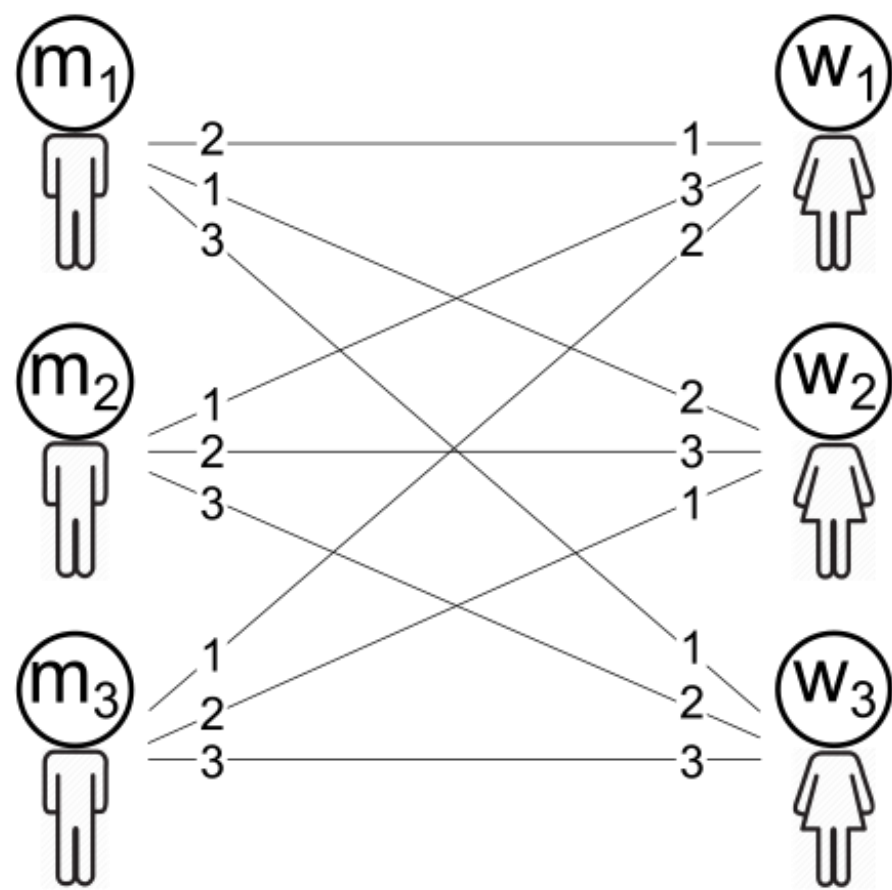
No stable matching procedure can be strategyproof.



Instance  $I_0$

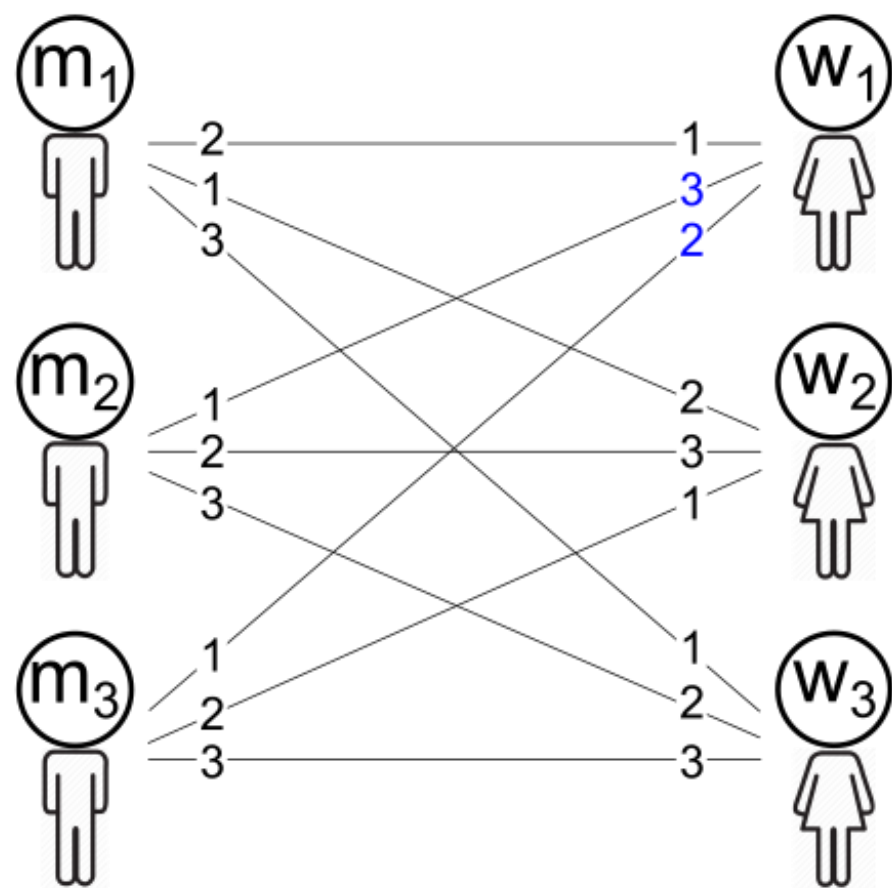


No stable matching procedure can be strategyproof.



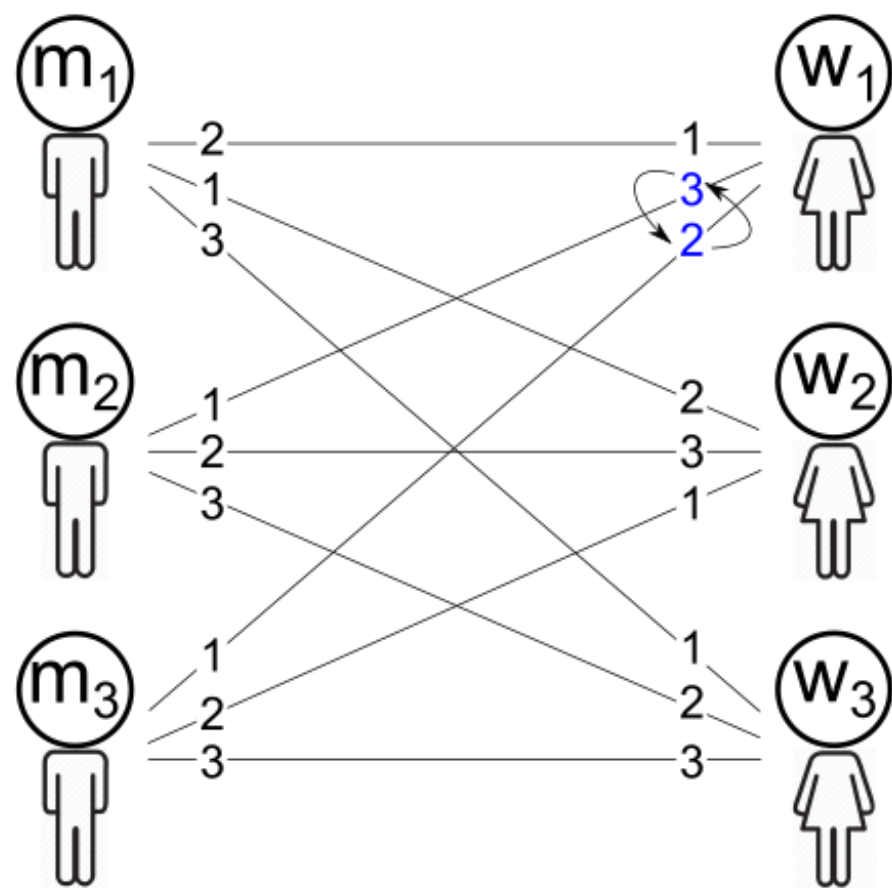
Instance  $I_0$

No stable matching procedure can be strategyproof.



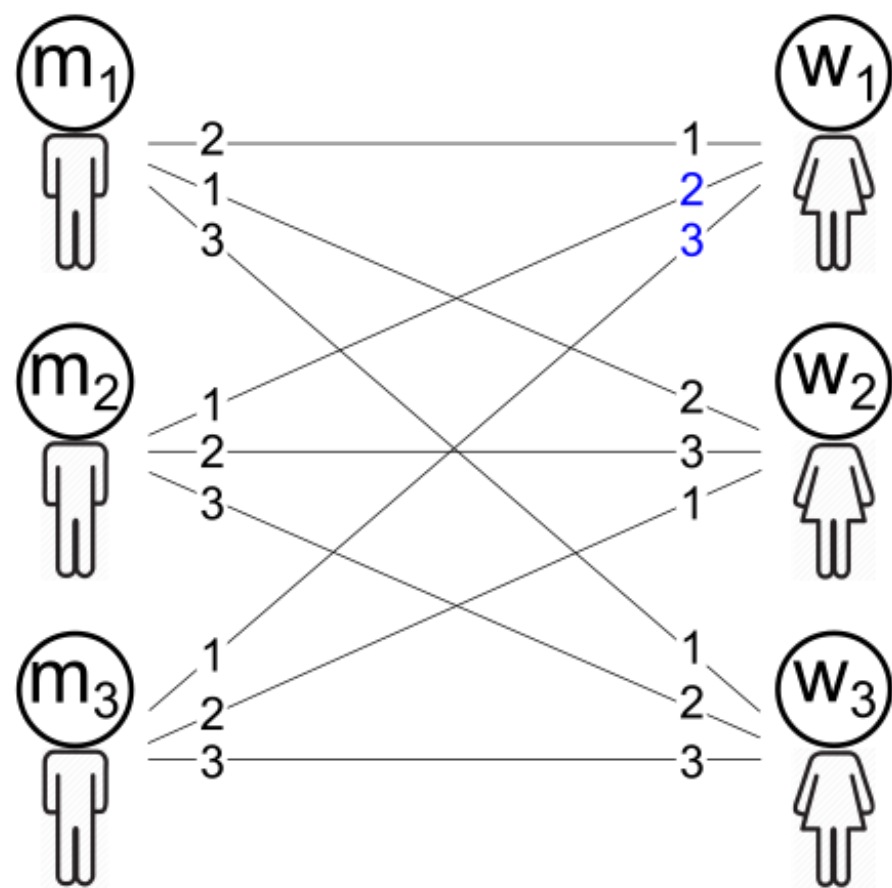
Instance  $I_0$

No stable matching procedure can be strategyproof.



Instance  $I_0$

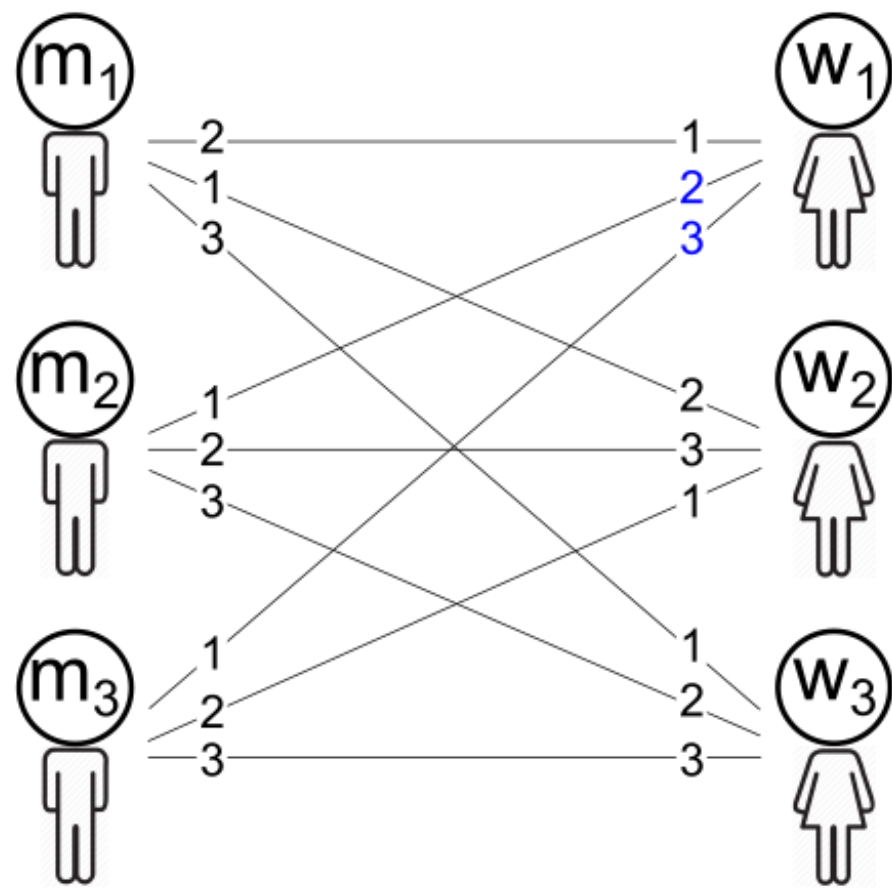
No stable matching procedure can be strategyproof.



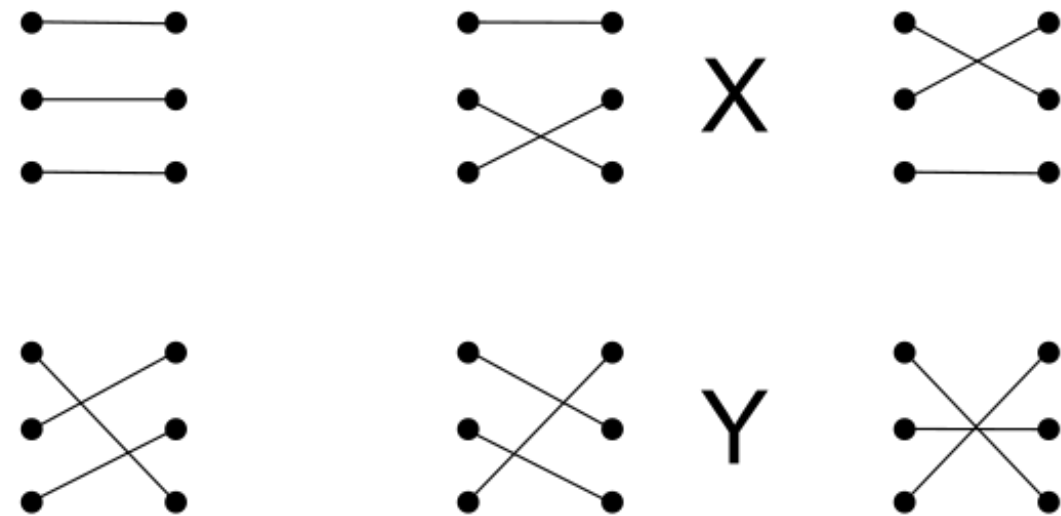
Instance  $I_1$



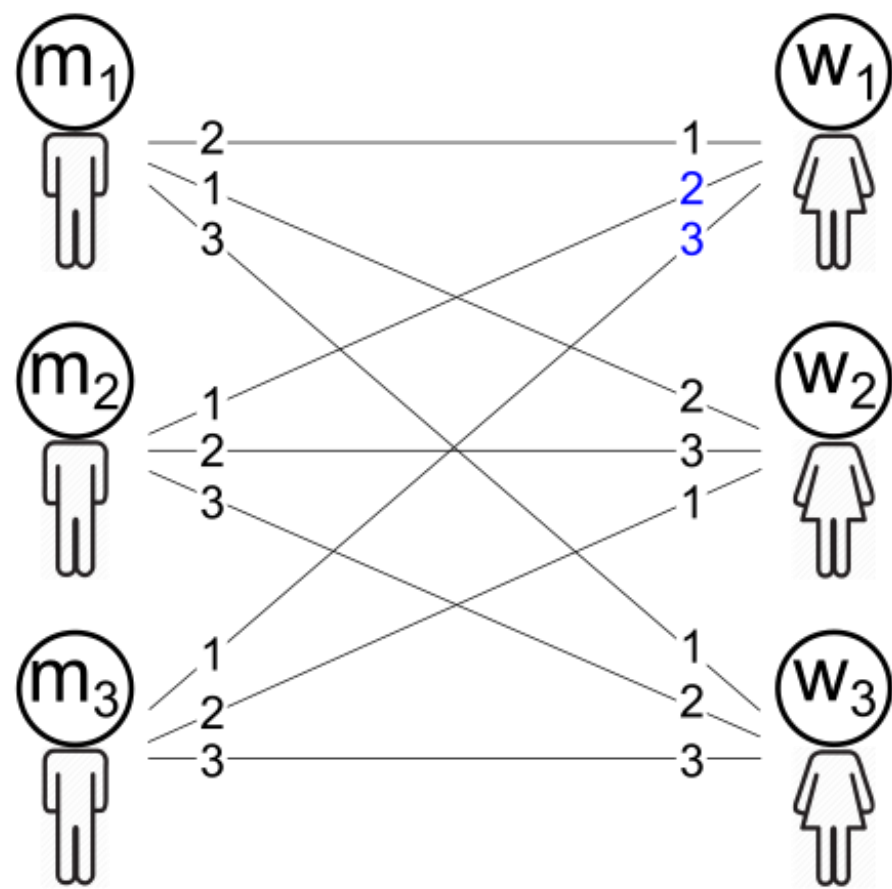
No stable matching procedure can be strategyproof.



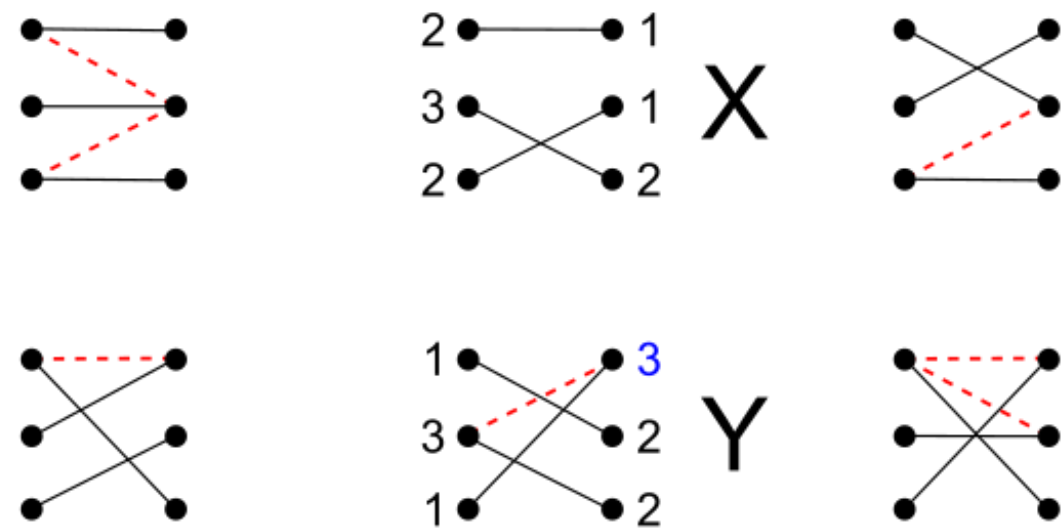
Instance  $I_1$



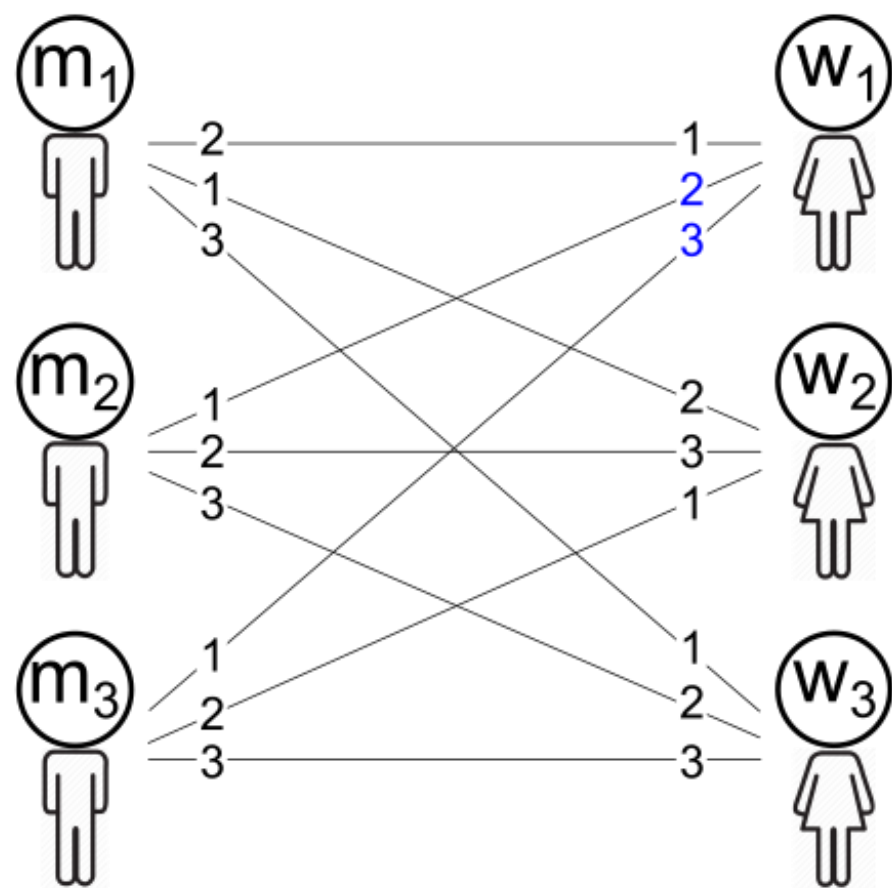
No stable matching procedure can be strategyproof.



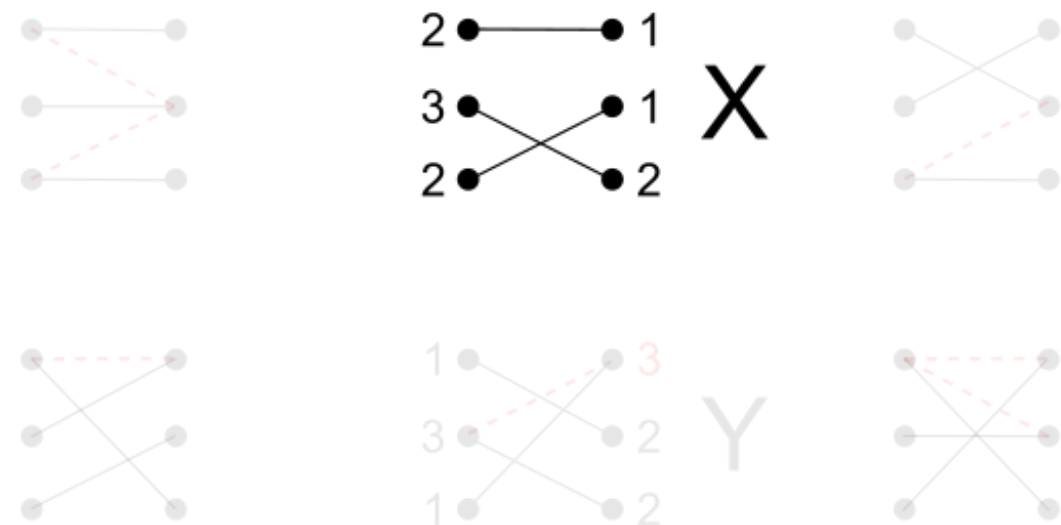
Instance  $I_1$



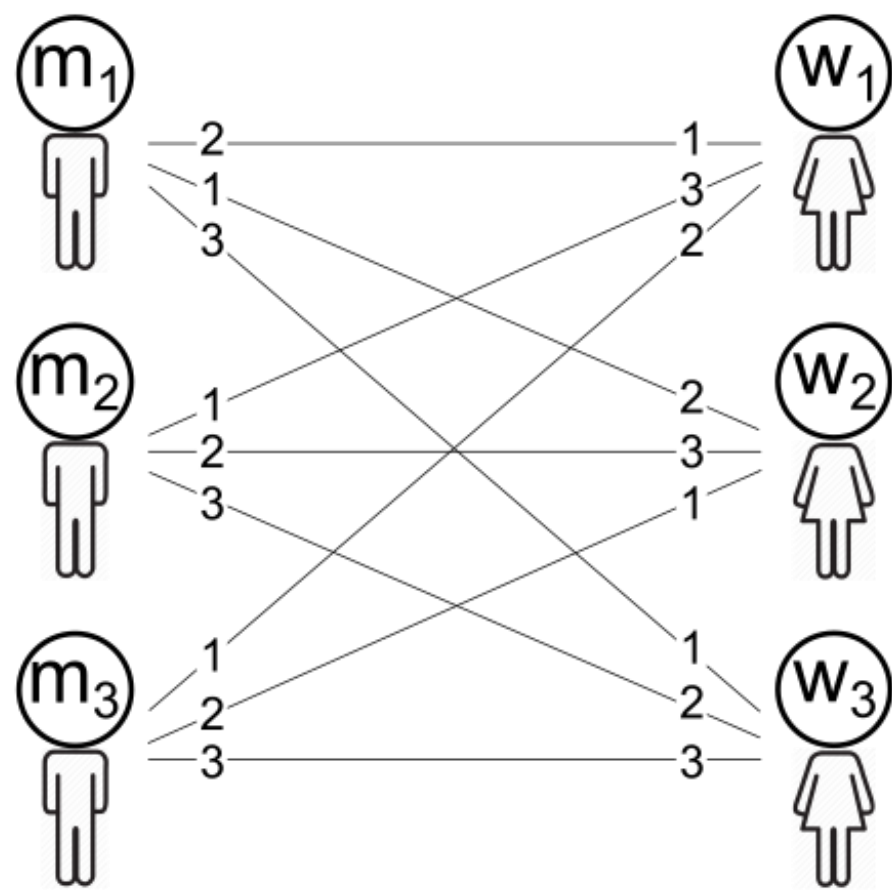
No stable matching procedure can be strategyproof.



Instance  $I_1$

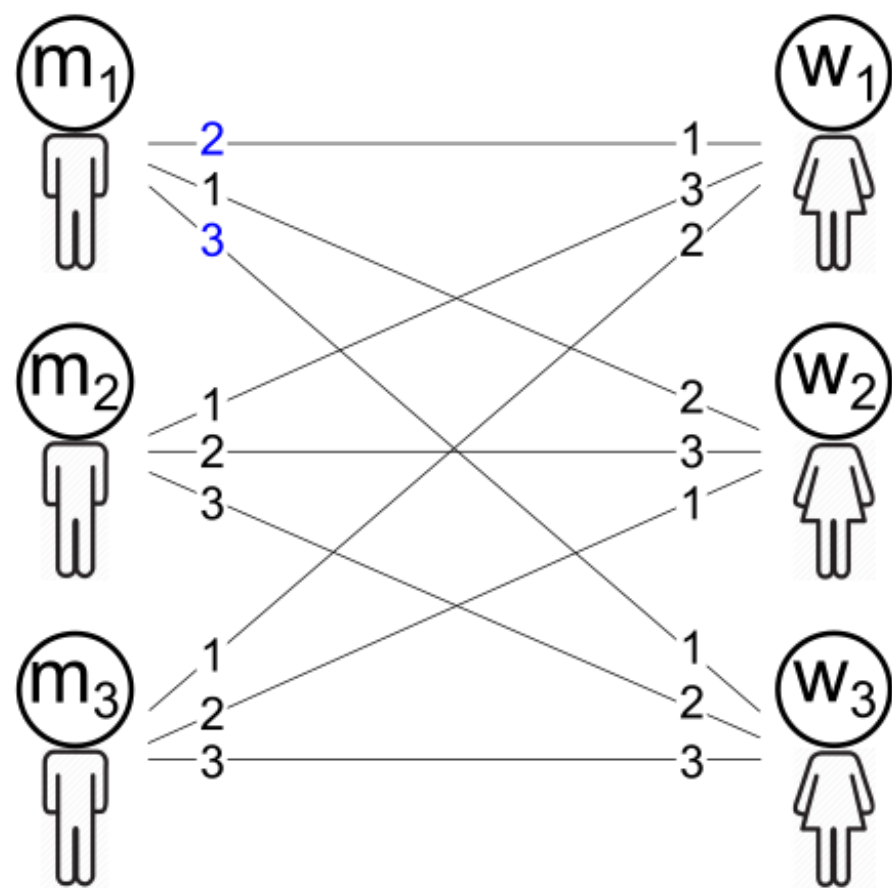


No stable matching procedure can be strategyproof.



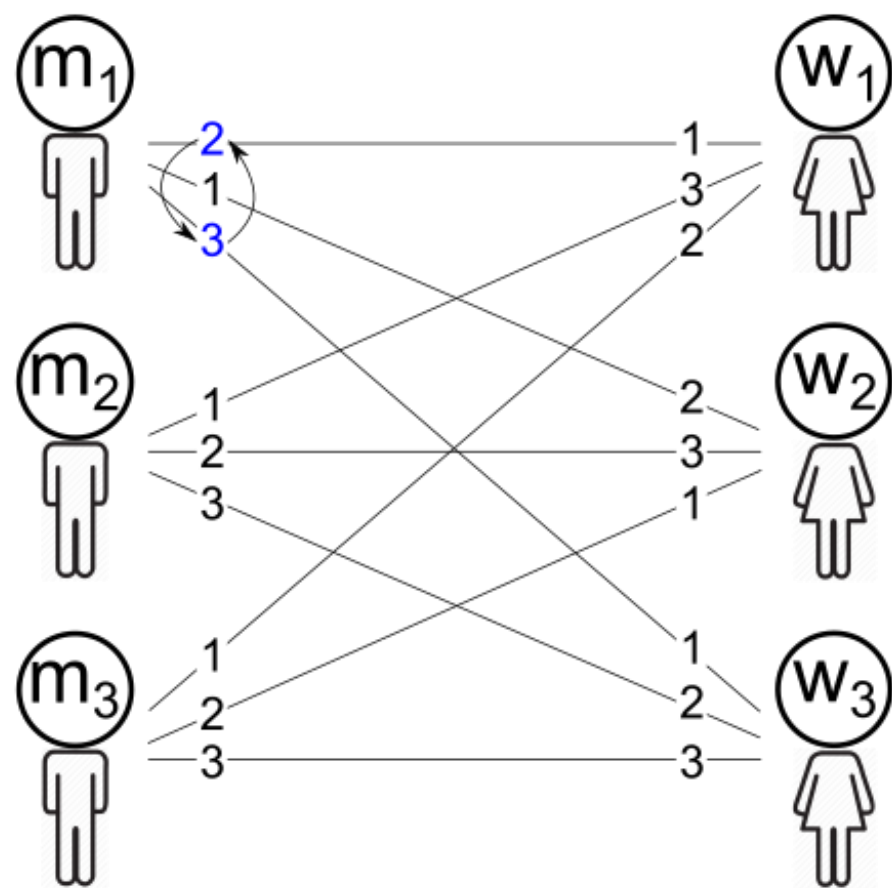
Instance  $I_0$

No stable matching procedure can be strategyproof.



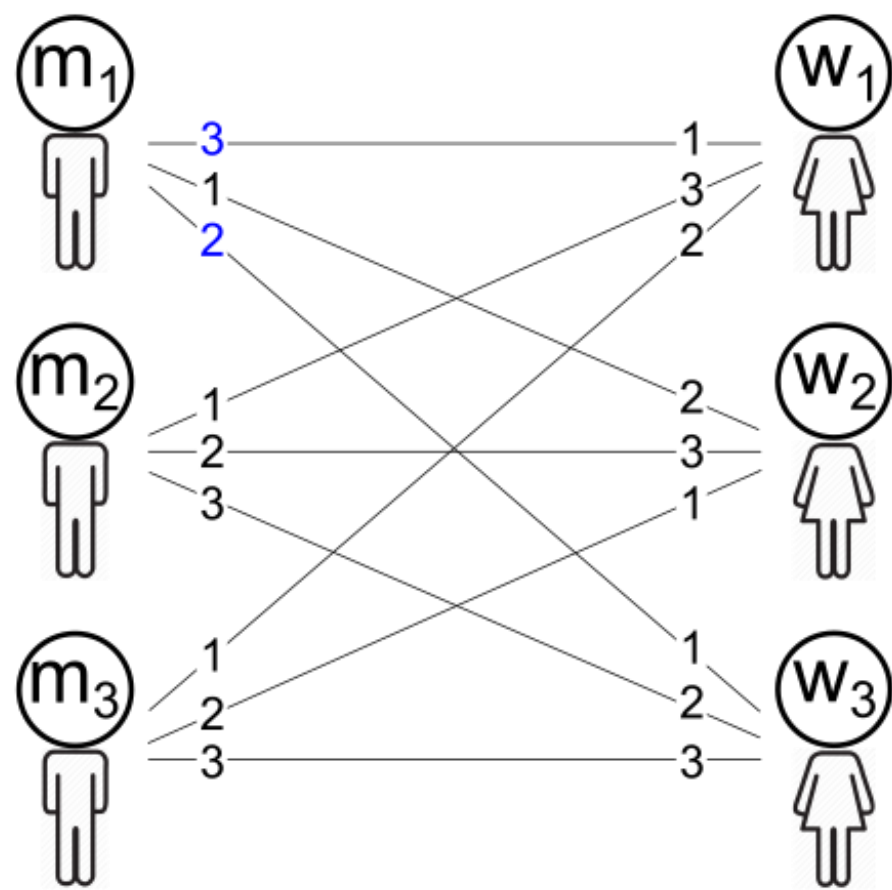
Instance  $I_0$

No stable matching procedure can be strategyproof.



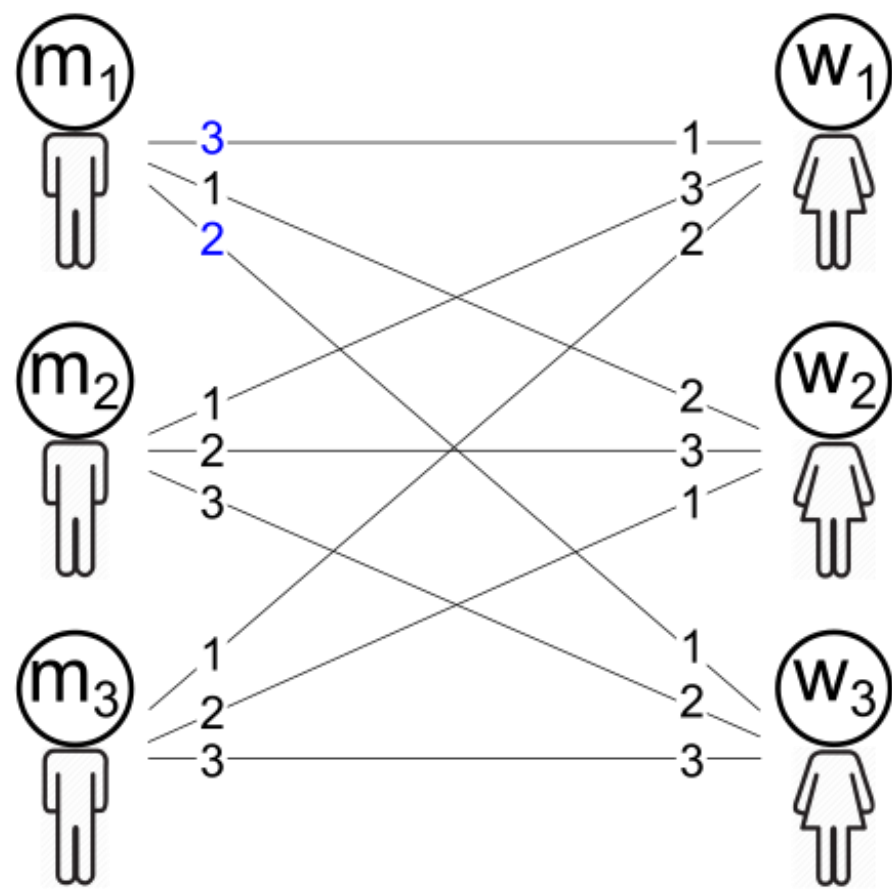
Instance  $I_0$

No stable matching procedure can be strategyproof.

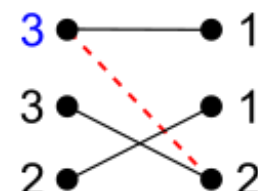
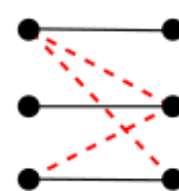


Instance  $I_2$

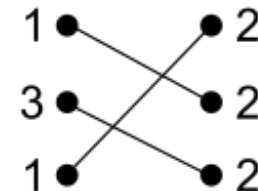
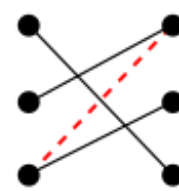
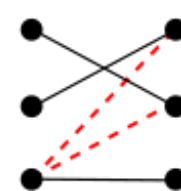
No stable matching procedure can be strategyproof.



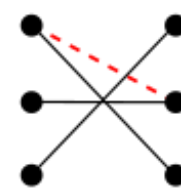
Instance  $I_2$



X

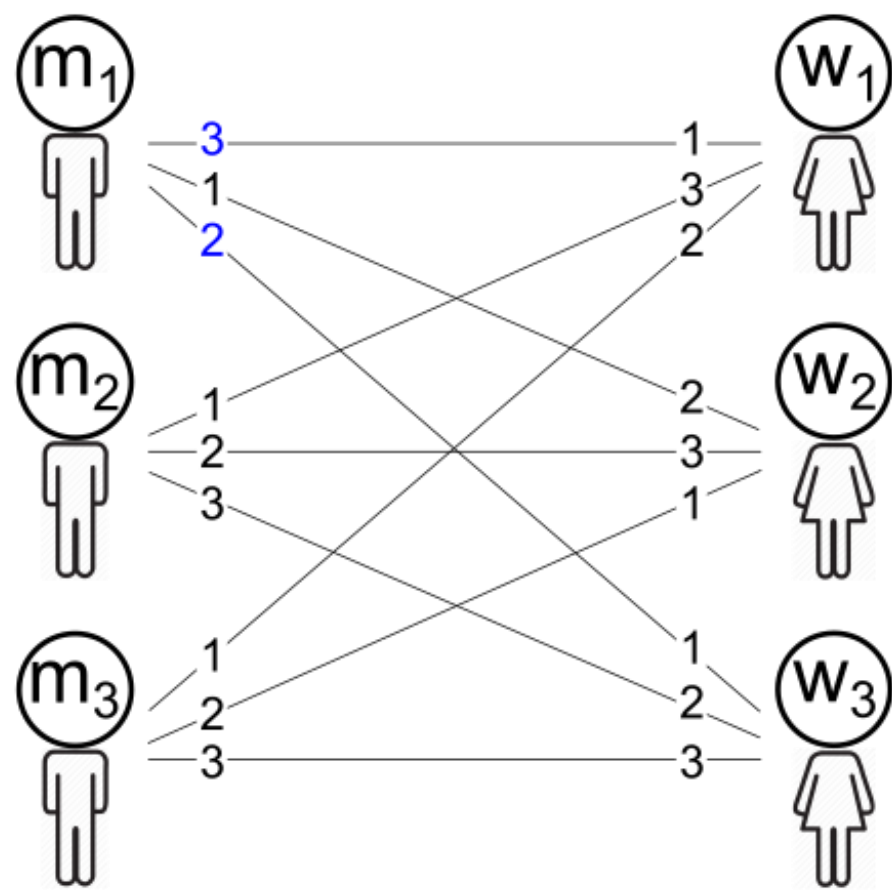


Y

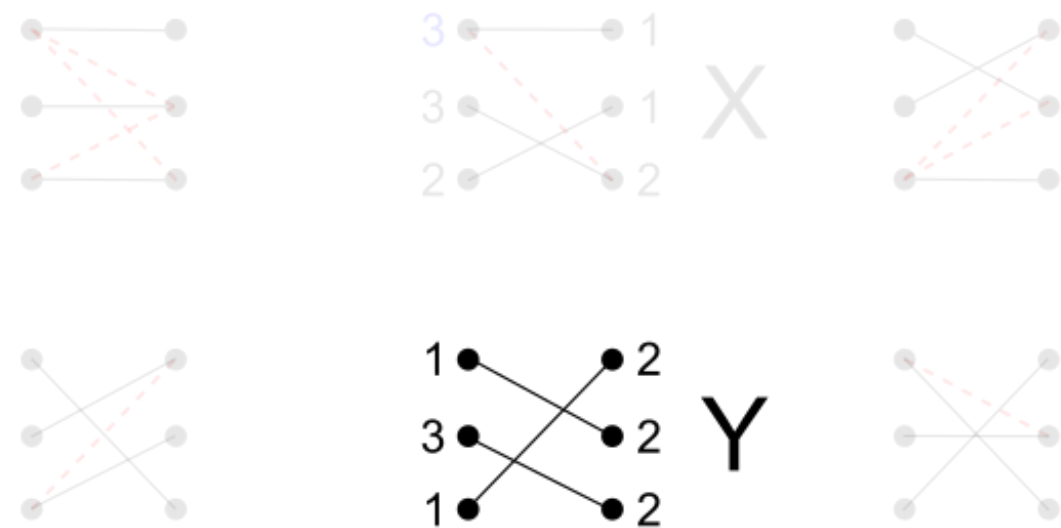




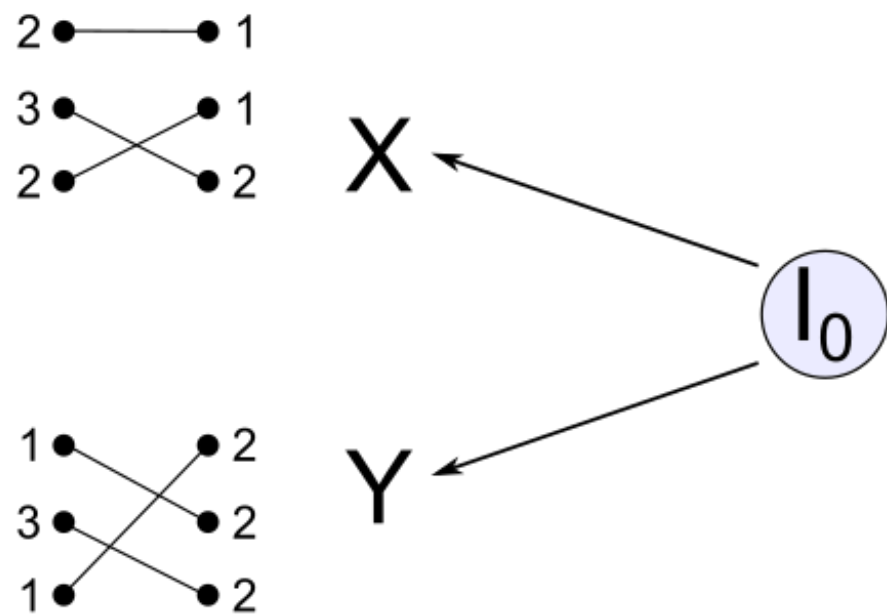
No stable matching procedure can be strategyproof.



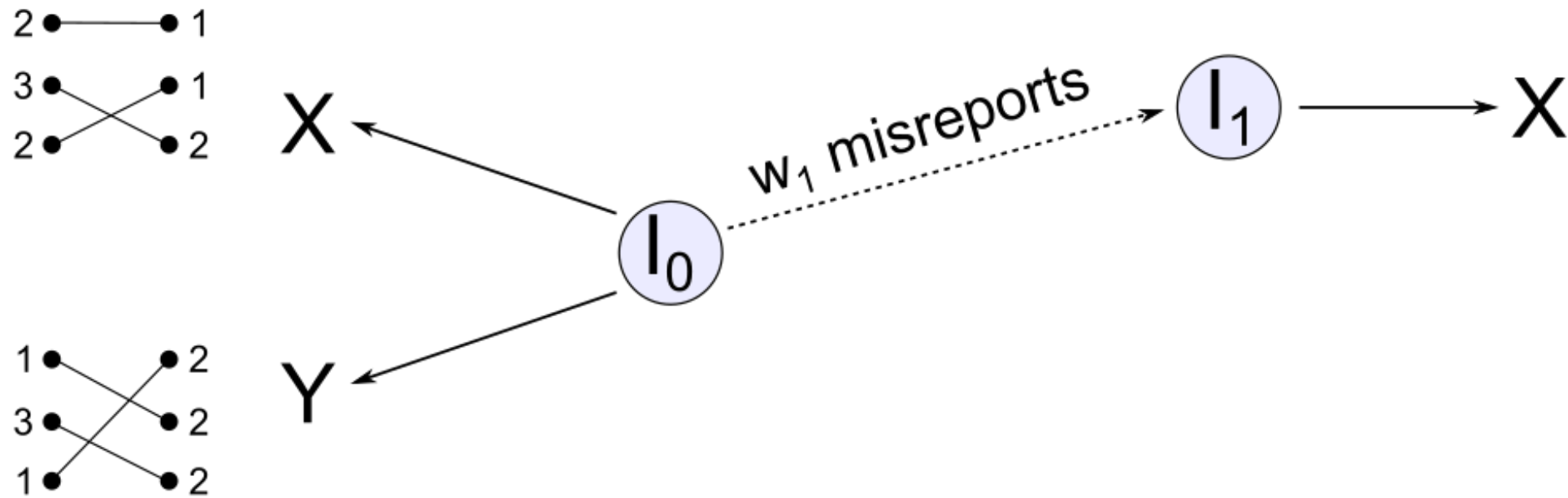
Instance  $I_2$



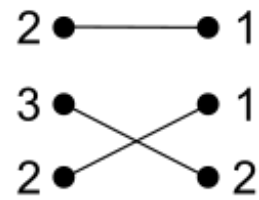
No stable matching procedure can be strategyproof.



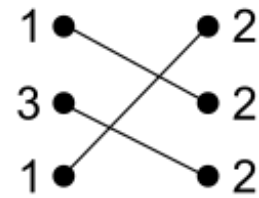
No stable matching procedure can be strategyproof.



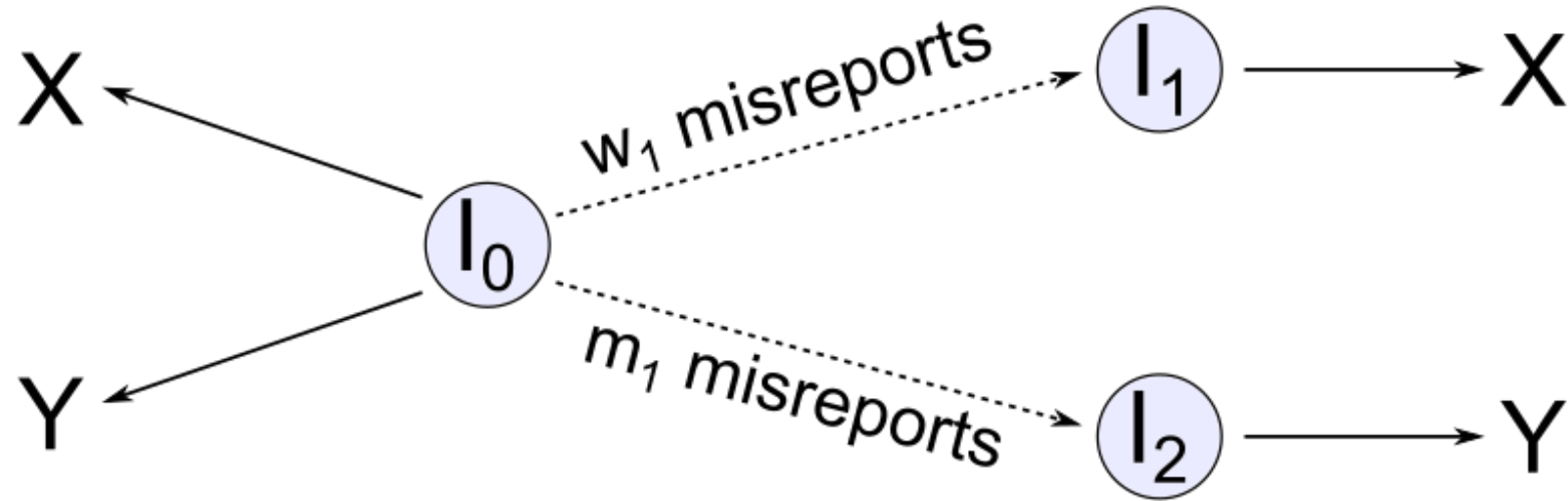
No stable matching procedure can be strategyproof.



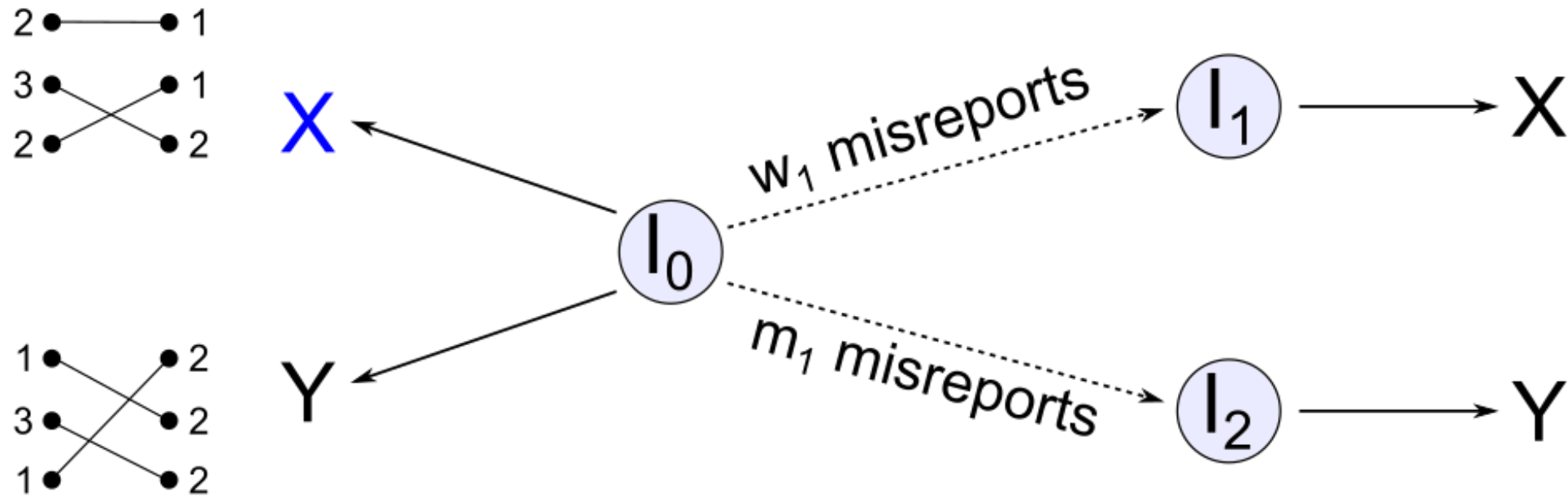
X



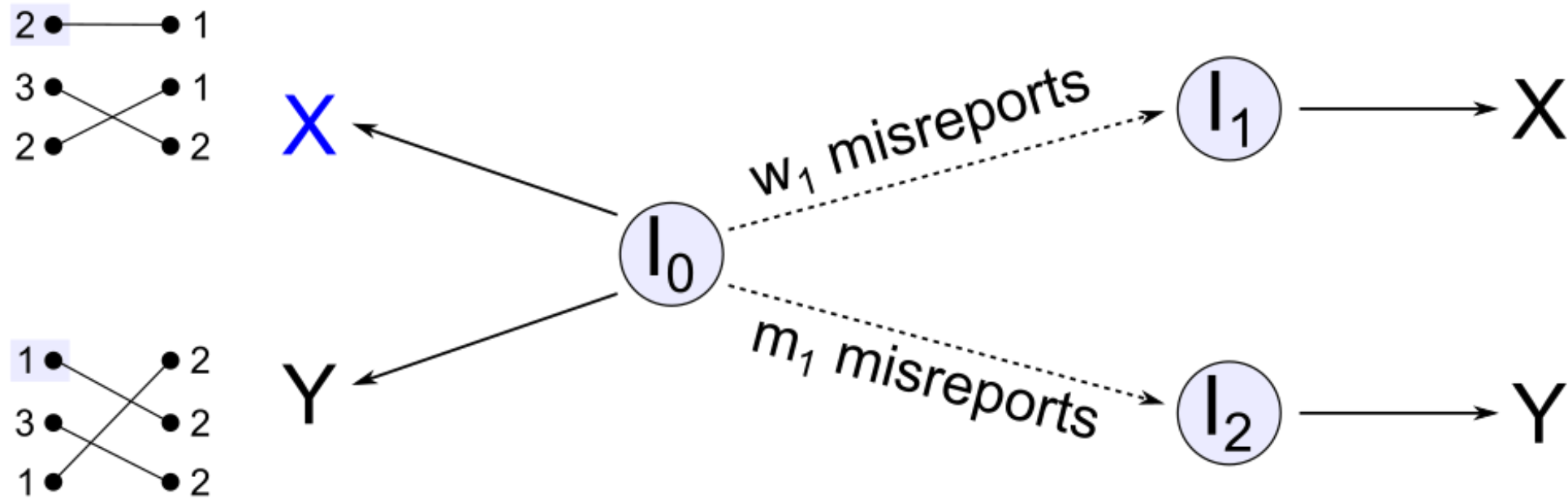
Y



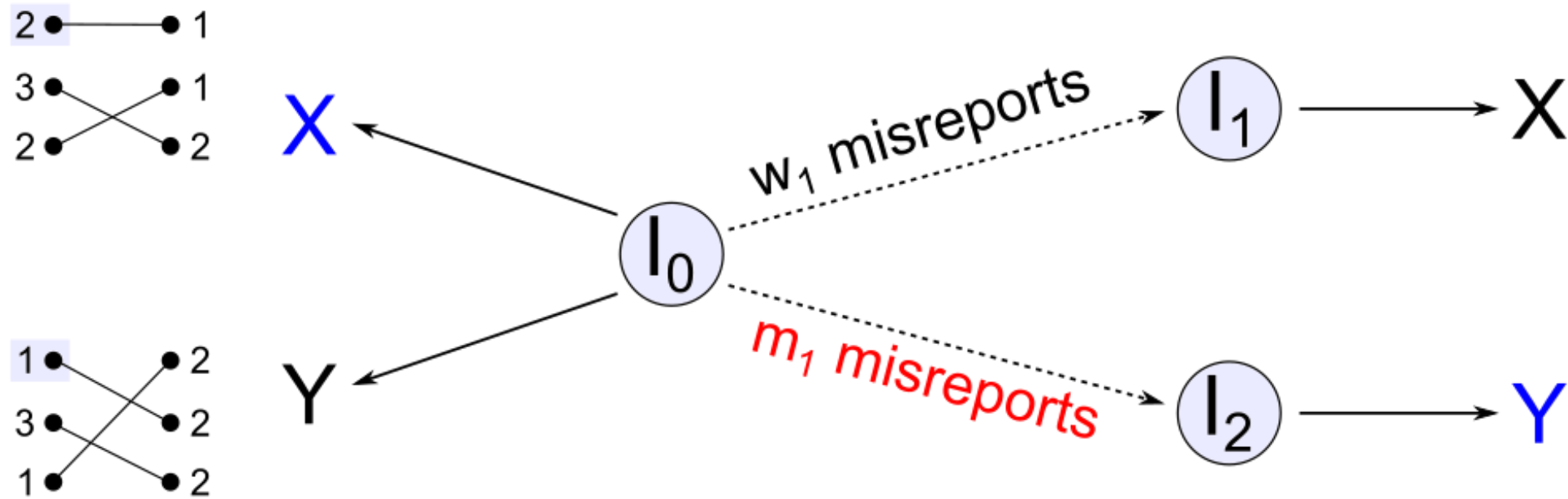
No stable matching procedure can be strategyproof.



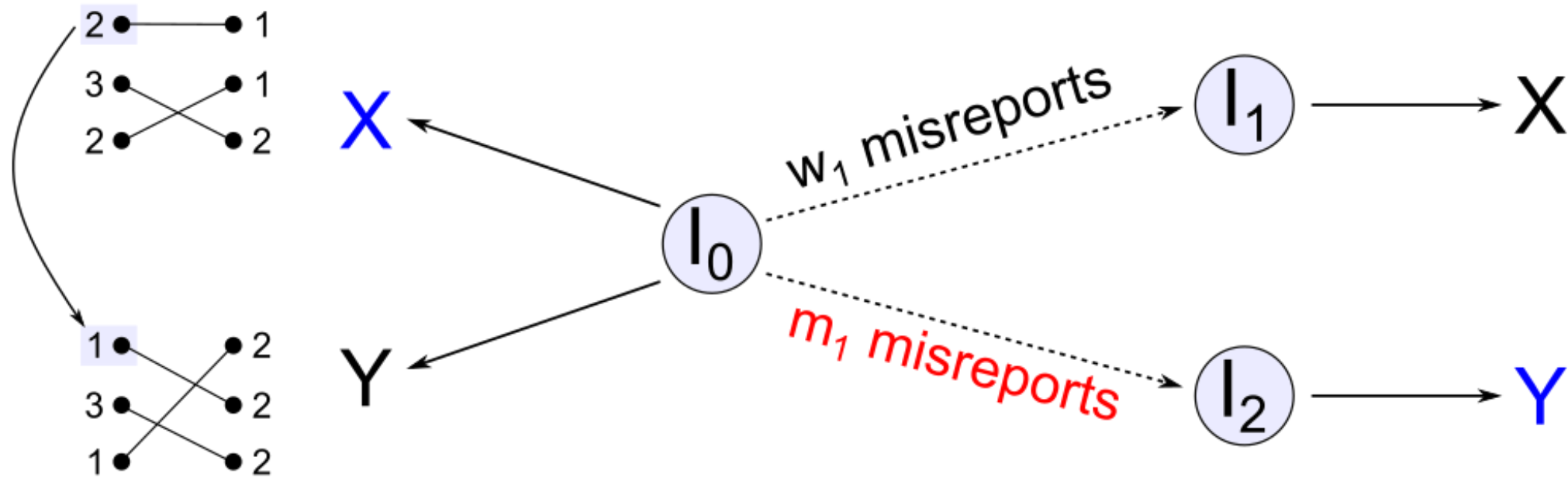
No stable matching procedure can be strategyproof.



No stable matching procedure can be strategyproof.

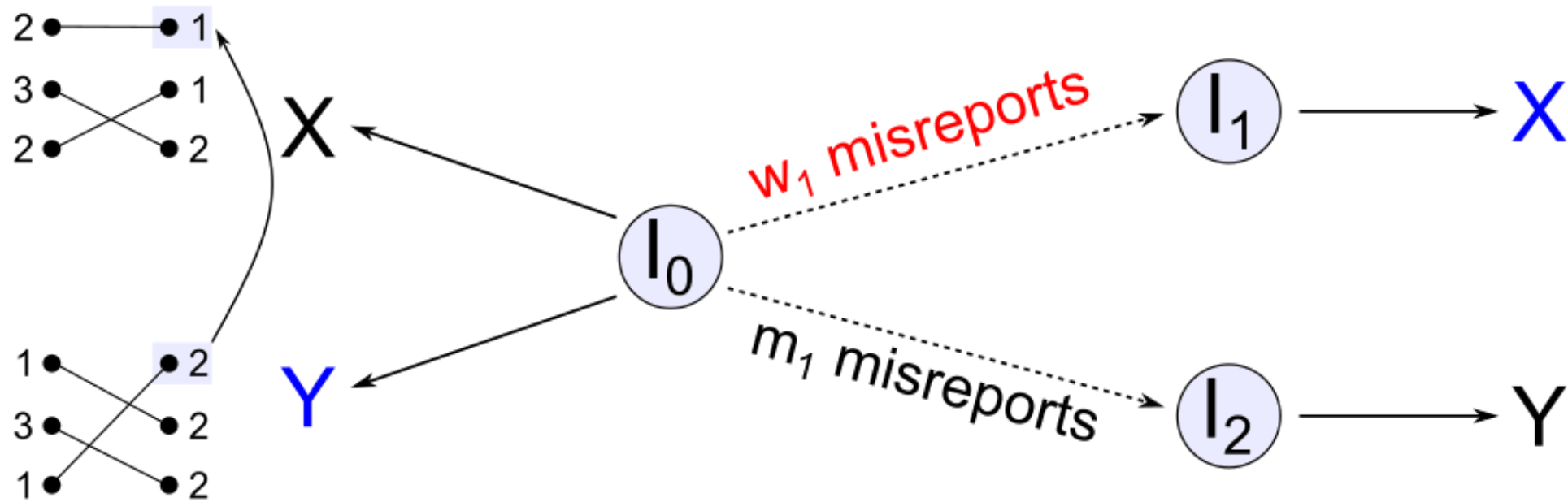


No stable matching procedure can be strategyproof.

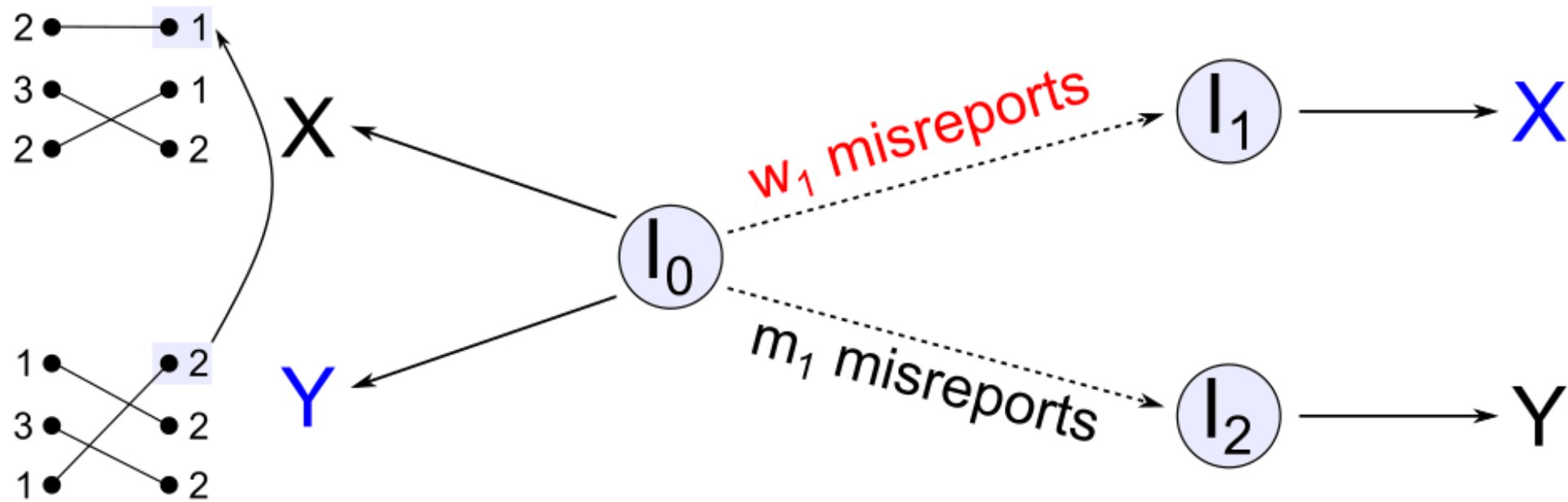




No stable matching procedure can be strategyproof.



No stable matching procedure can be strategyproof.



# RECAP

# RECAP

DA is strategyproof for the proposing side (men)  
but can be manipulated by the proposed-to side (women).

# RECAP

DA is strategyproof for the proposing side (men)  
but can be manipulated by the proposed-to side (women).

An optimal manipulation strategy is "inconspicuous" w/o loss of generality  
and thus can be efficiently computed.

# RECAP

DA is strategyproof for the proposing side (men)  
but can be manipulated by the proposed-to side (women).

An optimal manipulation strategy is "inconspicuous" w/o loss of generality  
and thus can be efficiently computed.

Optimal manipulation is stability-preserving (w.r.t. true preferences).

# RECAP

DA is strategyproof for the proposing side (men)  
but can be manipulated by the proposed-to side (women).

An optimal manipulation strategy is "inconspicuous" w/o loss of generality  
and thus can be efficiently computed.

Optimal manipulation is stability-preserving (w.r.t. true preferences).

No stable matching procedure is strategyproof for all agents.

# Next Time

## Finding Fair Stable Matchings





# References

- DA algorithm fails to be strategyproof.

Lester Dubins and David Freedman

"Machiavelli and the Gale-Shapley Algorithm"

*American Mathematical Monthly*, 88(7), 1981 pg 485-494

- No stable matching procedure is strategyproof.

Alvin E Roth

"The Economics of Matching: Stability and Incentives"

*Mathematics of Operations Research*, 7(4), 1982 pg 617-628

# References

- An algorithm for finding an optimal manipulation for a woman.

Chung-Piaw Teo, Jay Sethuraman, and Wee-Peng Tan

“Gale-Shapley Stable Marriage Problem Revisited: Strategic Issues and Applications”

*Management Science*, 47(9), 2001 pg 252–1267

- Optimally manipulated marriages are stable.

Rohit Vaish and Dinesh Garg

“Manipulating Gale-Shapley Algorithm: Preserving Stability and Remaining Inconspicuous”

*IJCAI 2017*, pg 437-443



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is strategyproof for the men.

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

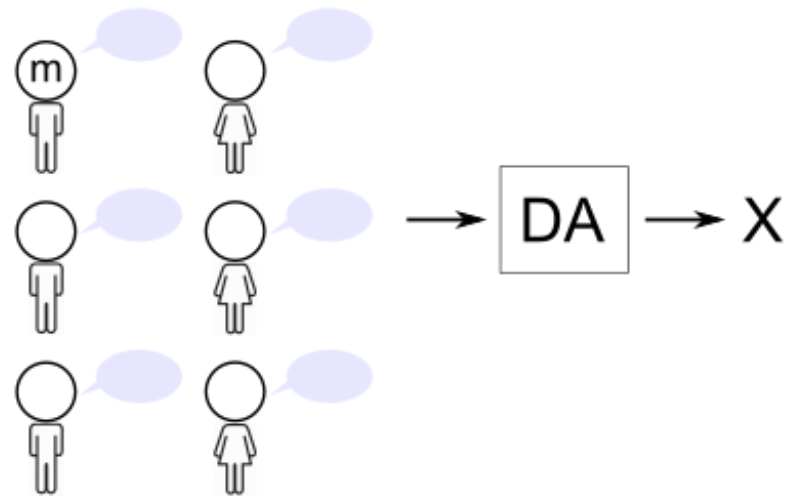
Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

True profile  $P = (P_{-m}, P_m)$

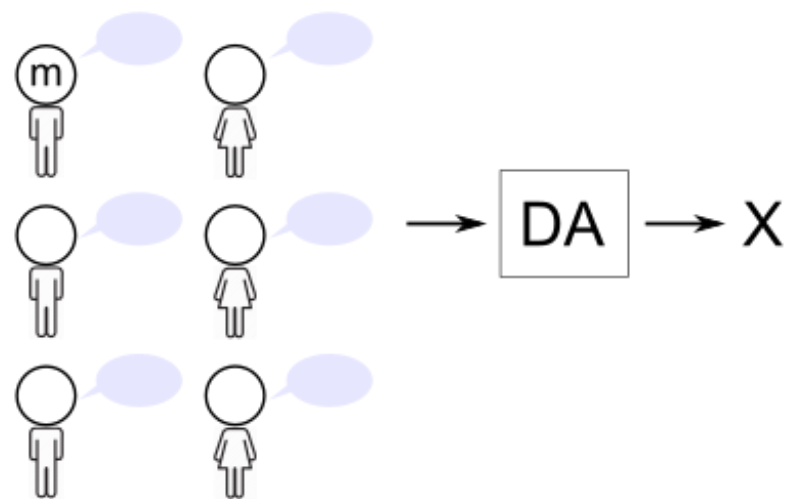


[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

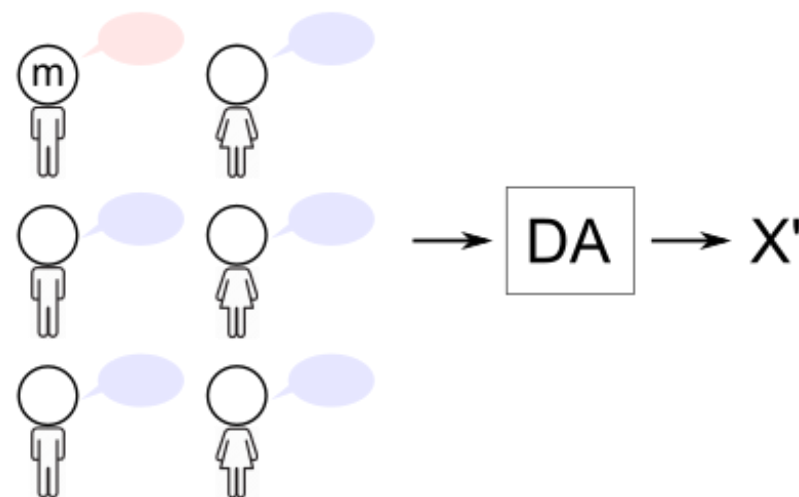
**DA algorithm is strategyproof for the men.**

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

True profile  $P = (P_{-m}, P_m)$



Manipulated profile  $P' = (P_{-m}, P'_m)$

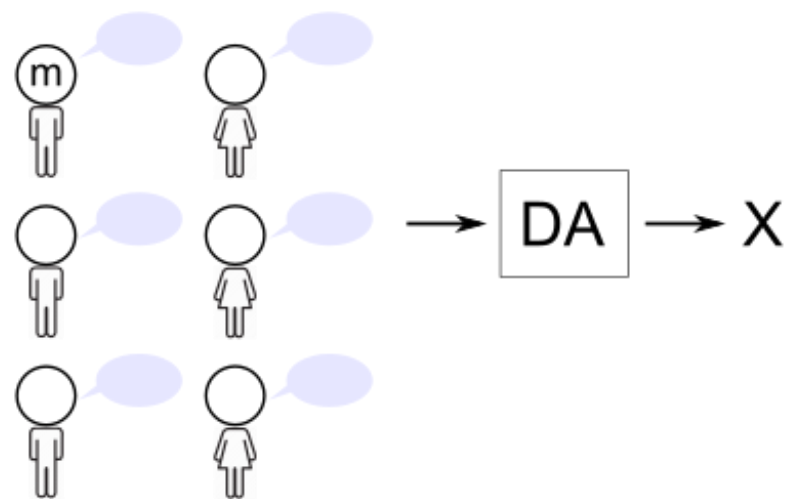


[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

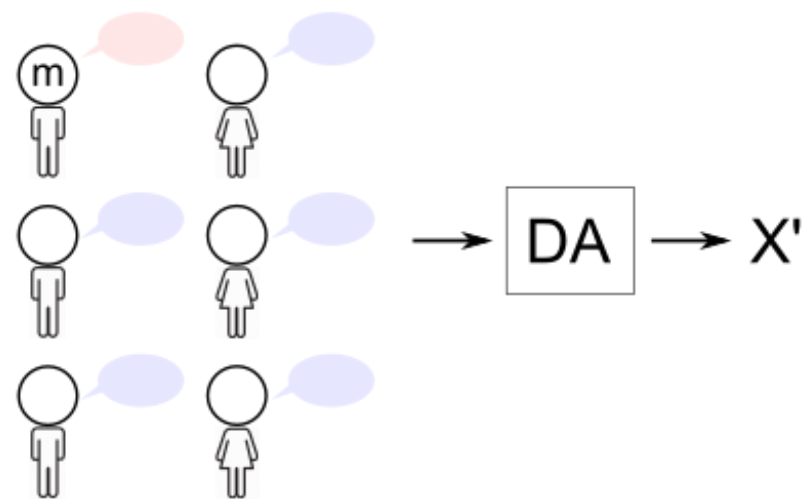
## DA algorithm is strategyproof for the men.

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

True profile  $P = (P_{-m}, P_m)$



Manipulated profile  $P' = (P_{-m}, P'_m)$



$P_m \quad \dots > X'(m) > \dots > X(m) > \dots$

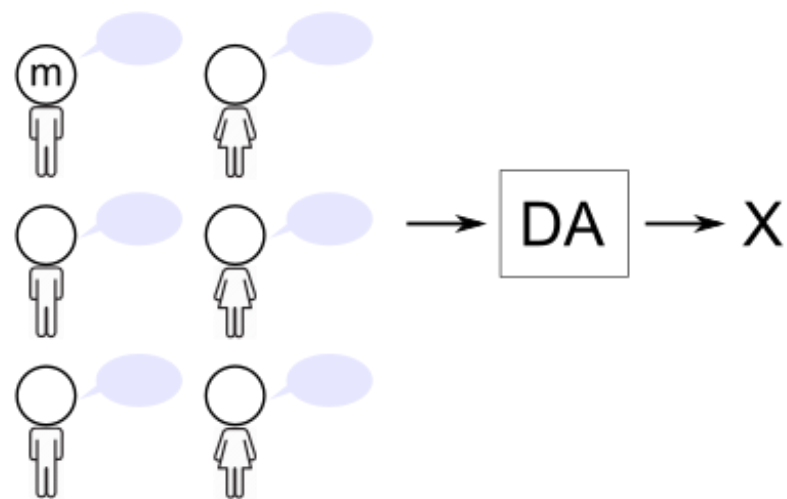


[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

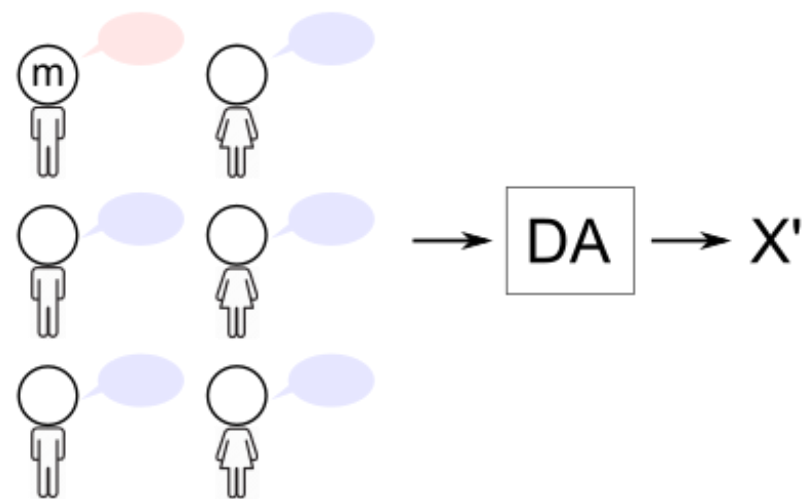
## DA algorithm is strategyproof for the men.

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

True profile  $P = (P_{-m}, P_m)$



Manipulated profile  $P' = (P_{-m}, P'_m)$



$P_m \quad \dots > X'(m) > \dots > X(m) > \dots$

We will use three lemmas to derive a contradiction.

1

## Simplicity

If there is a feasible strategy for manipulation, then there is a "simple" feasible strategy that achieves the same outcome.

## Simplicity

1

If there is a feasible strategy for manipulation, then there is a "simple" feasible strategy that achieves the same outcome.

## Brotherhood

2

All men are weakly better off under  $P'$  (compared to  $P$ ).

## Simplicity

1

If there is a feasible strategy for manipulation, then there is a "simple" feasible strategy that achieves the same outcome.

## Brotherhood

2

All men are weakly better off under  $P'$  (compared to  $P$ ).

## No new proposal

3

If a man proposes to a woman during DA on the profile  $P'$ , then he must also propose to her during DA on the profile  $P$ .

# 1

## Simplicity

Let  $X' = DA(P')$ . Then,  $m$  is matched to  $X'(m)$  under DA on the profile  $P'' = (P_{-m}, P''_m)$  obtained from his true list  $P_m$  by promoting  $X'(m)$  to the top.

# 1

## Simplicity

Let  $X' = DA(P')$ . Then,  $m$  is matched to  $X'(m)$  under  $DA$  on the profile  $P'' = (P_{-m}, P''_m)$  obtained from his true list  $P_m$  by promoting  $X'(m)$  to the top.

The matching  $X'$  is stable w.r.t. the profile  $P'$ .

# 1

## Simplicity

Let  $X' = DA(P')$ . Then,  $m$  is matched to  $X'(m)$  under  $DA$  on the profile  $P'' = (P_{-m}, P''_m)$  obtained from his true list  $P_m$  by promoting  $X'(m)$  to the top.

The matching  $X'$  is stable w.r.t. the profile  $P'$ .

Then,  $X'$  must also be stable w.r.t. the profile  $P''$ .

# 1

## Simplicity

Let  $X' = DA(P')$ . Then,  $m$  is matched to  $X'(m)$  under  $DA$  on the profile  $P'' = (P_{-m}, P''_m)$  obtained from his true list  $P_m$  by promoting  $X'(m)$  to the top.

The matching  $X'$  is stable w.r.t. the profile  $P'$ .

Then,  $X'$  must also be stable w.r.t. the profile  $P''$ .

(The manipulator  $m$  gets his top choice in  $P''$  and therefore doesn't block  $X'$ . Any other blocking pair must also block  $X'$  w.r.t.  $P'$ , but that would contradict stability of  $DA$  algorithm.)



# 1

## Simplicity

Let  $X' = DA(P')$ . Then,  $m$  is matched to  $X'(m)$  under DA on the profile  $P'' = (P_{-m}, P''_m)$  obtained from his true list  $P_m$  by promoting  $X'(m)$  to the top.

The matching  $X'$  is stable w.r.t. the profile  $P'$ .

Then,  $X'$  must also be stable w.r.t. the profile  $P''$ .

(The manipulator  $m$  gets his top choice in  $P''$  and therefore doesn't block  $X'$ . Any other blocking pair must also block  $X'$  w.r.t.  $P'$ , but that would contradict stability of DA algorithm.)

When DA is run on  $P''$ , we get the men-optimal stable matching, say  $X''$ , w.r.t.  $P''$ .

# 1

## Simplicity

Let  $X' = DA(P')$ . Then,  $m$  is matched to  $X'(m)$  under DA on the profile  $P'' = (P_{-m}, P''_m)$  obtained from his true list  $P_m$  by promoting  $X'(m)$  to the top.

The matching  $X'$  is stable w.r.t. the profile  $P'$ .

Then,  $X'$  must also be stable w.r.t. the profile  $P''$ .

(The manipulator  $m$  gets his top choice in  $P''$  and therefore doesn't block  $X'$ . Any other blocking pair must also block  $X'$  w.r.t.  $P'$ , but that would contradict stability of DA algorithm.)

When DA is run on  $P''$ , we get the men-optimal stable matching, say  $X''$ , w.r.t.  $P''$ .

Man  $m$  must weakly prefer  $X''(m)$  over  $X'(m)$  according to  $P''_m$ .

# 1

## Simplicity

Let  $X' = DA(P')$ . Then,  $m$  is matched to  $X'(m)$  under DA on the profile  $P'' = (P_{-m}, P''_m)$  obtained from his true list  $P_m$  by promoting  $X'(m)$  to the top.

The matching  $X'$  is stable w.r.t. the profile  $P'$ .

Then,  $X'$  must also be stable w.r.t. the profile  $P''$ .

(The manipulator  $m$  gets his top choice in  $P''$  and therefore doesn't block  $X'$ . Any other blocking pair must also block  $X'$  w.r.t.  $P'$ , but that would contradict stability of DA algorithm.)

When DA is run on  $P''$ , we get the men-optimal stable matching, say  $X''$ , w.r.t.  $P''$ .

Man  $m$  must weakly prefer  $X''(m)$  over  $X'(m)$  according to  $P''_m$ .

But the woman  $X'(m)$  is already his top choice in  $P''_m$ . So,  $X''(m) = X'(m)$ .

# 1

## Simplicity

Let  $X' = DA(P')$ . Then,  $m$  is matched to  $X'(m)$  under DA on the profile  $P'' = (P_{-m}, P''_m)$  obtained from his true list  $P_m$  by promoting  $X'(m)$  to the top.

The matching  $X'$  is stable w.r.t. the profile  $P'$ .

Then,  $X'$  must also be stable w.r.t. the profile  $P''$ .

(The manipulator  $m$  gets his top choice in  $P''$  and therefore doesn't block  $X'$ . Any other blocking pair must also block  $X'$  w.r.t.  $P'$ , but that would contradict stability of DA algorithm.)

When DA is run on  $P''$ , we get the men-optimal stable matching, say  $X''$ , w.r.t.  $P''$ .

Man  $m$  must weakly prefer  $X''(m)$  over  $X'(m)$  according to  $P''_m$ .

But the woman  $X'(m)$  is already his top choice in  $P''_m$ . So,  $X''(m) = X'(m)$ .



# 2

## Brotherhood

All men are weakly better off under  $X'$  (compared to  $X$ ).

# 2

## Brotherhood

All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

# 2

## Brotherhood

All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

# 2

## Brotherhood

All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .



# 2

## Brotherhood

All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .

# 2

## Brotherhood

All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .

$P$

$r$  —  $X(r)$

# 2

## Brotherhood

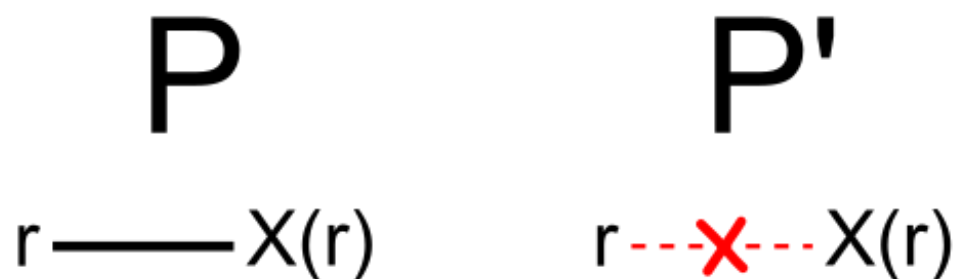
All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .



# 2

## Brotherhood

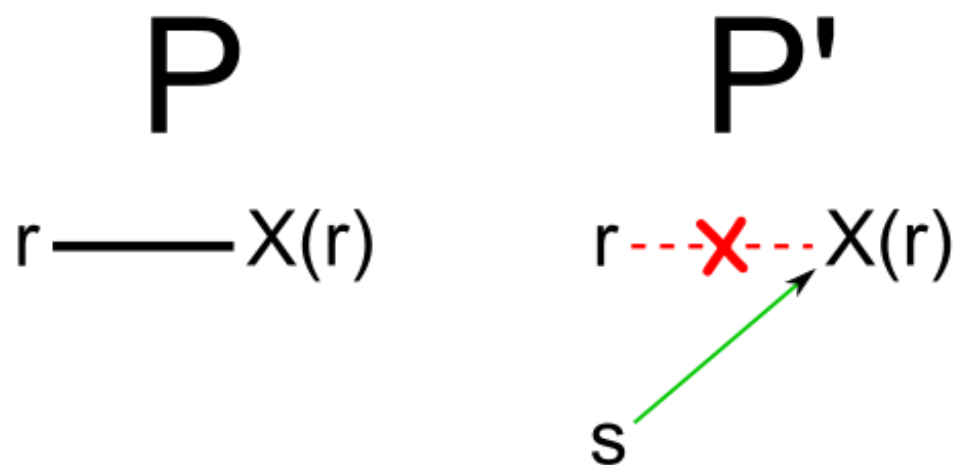
All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .



# 2

## Brotherhood

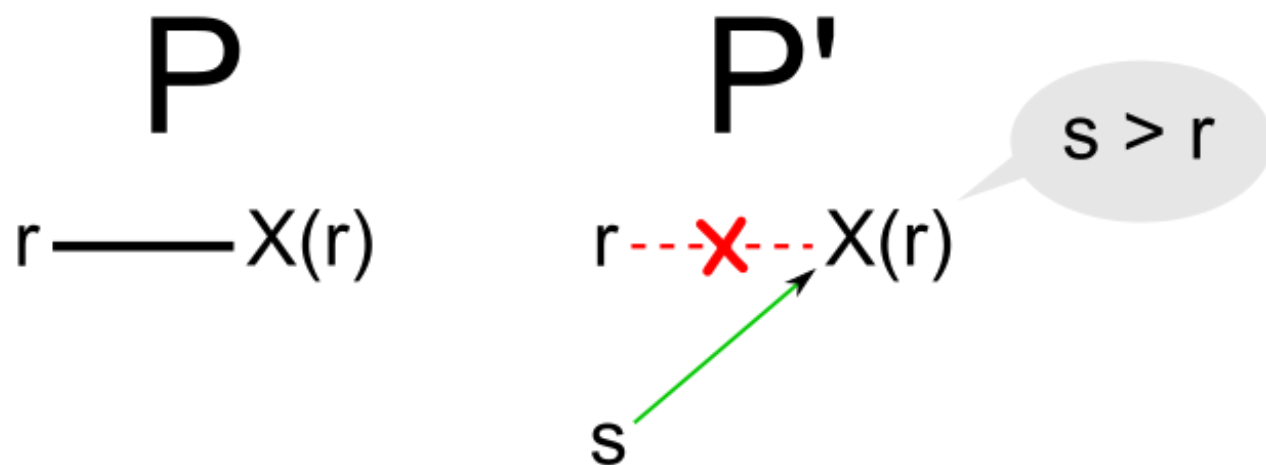
All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .



# 2

## Brotherhood

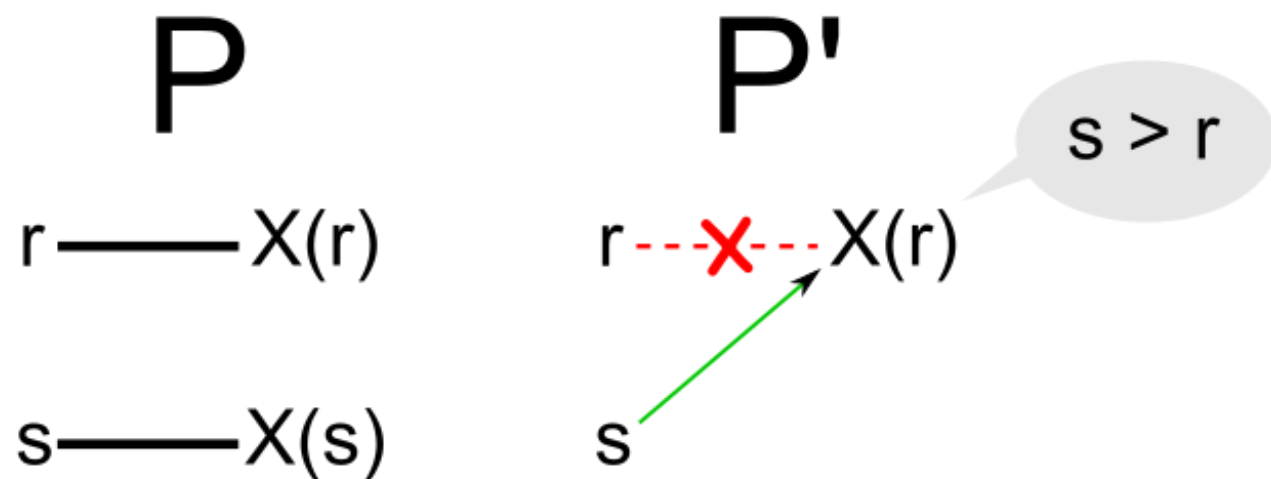
All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .



# 2

## Brotherhood

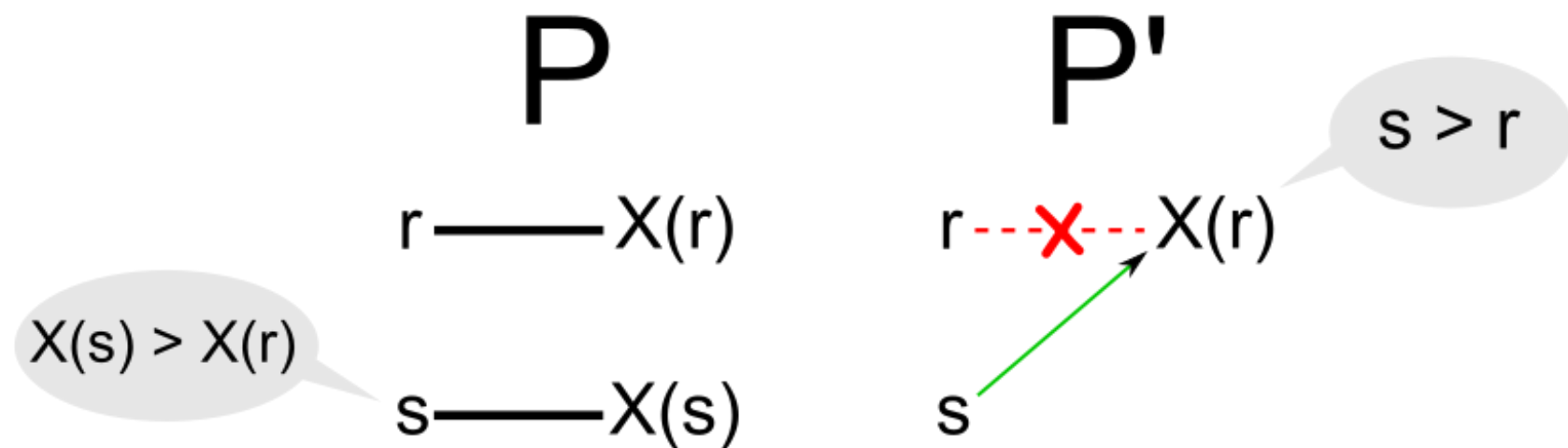
All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .



# 2

## Brotherhood

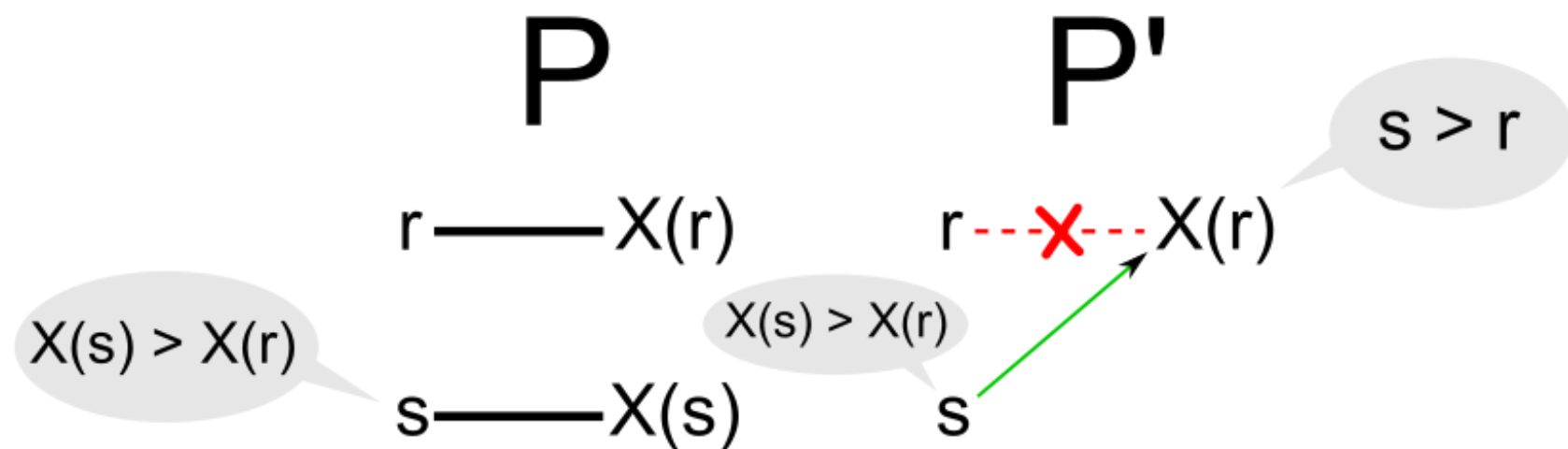
All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .





# 2

## Brotherhood

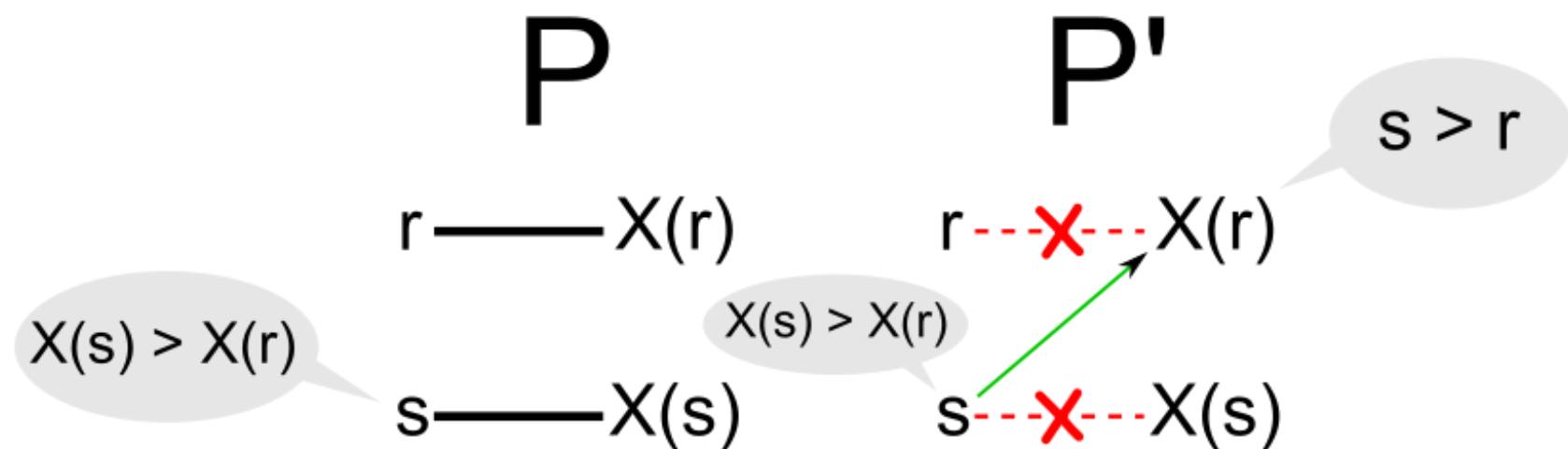
All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .



# 2

## Brotherhood

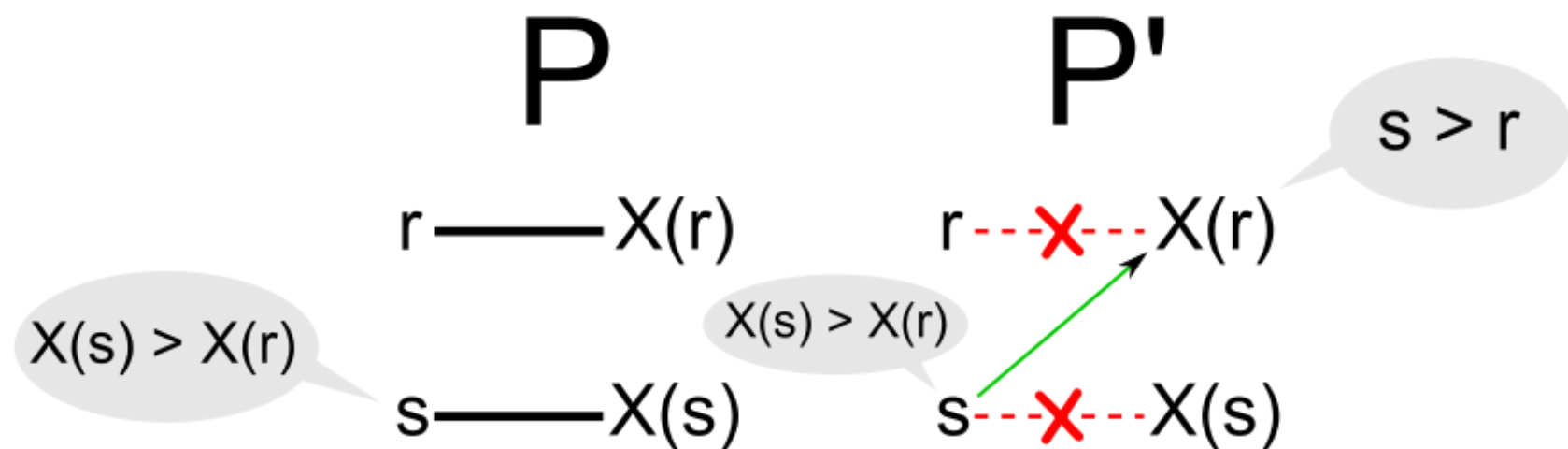
All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .



Then,  $s$  must have been rejected by  $X(s)$  prior to round  $k$ ---a contradiction.

# 2

## Brotherhood

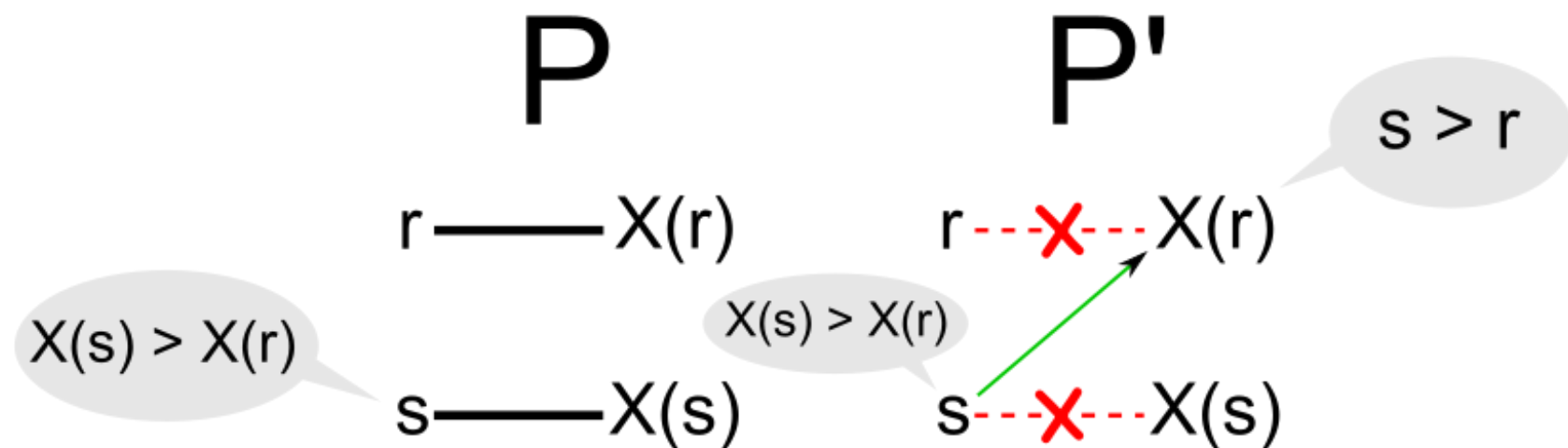
All men are weakly better off under  $X'$  (compared to  $X$ ).

Suppose, for contradiction, that some man is worse off under  $X'$ .

All such men must be truthful.

Among such men, let  $r$  be the *earliest* to be rejected by his  $X$ -partner during  $DA(P')$ .

Suppose  $X(r)$  rejects  $r$  in favor of the man  $s$  in round  $k$ .



Then,  $s$  must have been rejected by  $X(s)$  prior to round  $k$ ---a contradiction. ■

# 3

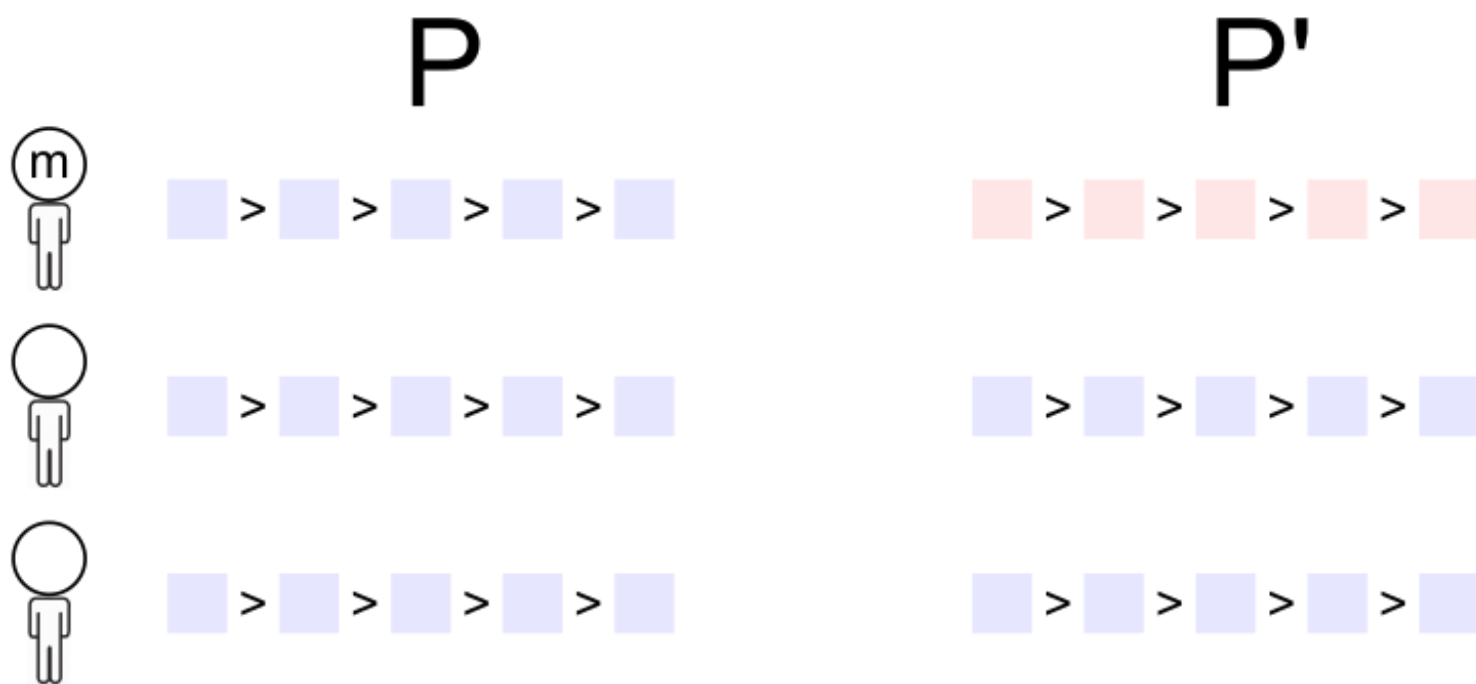
## No new proposal

If a man proposes to a woman during DA on the profile  $P'$ , then he must also propose to her during DA on the profile  $P$ .

# 3

## No new proposal

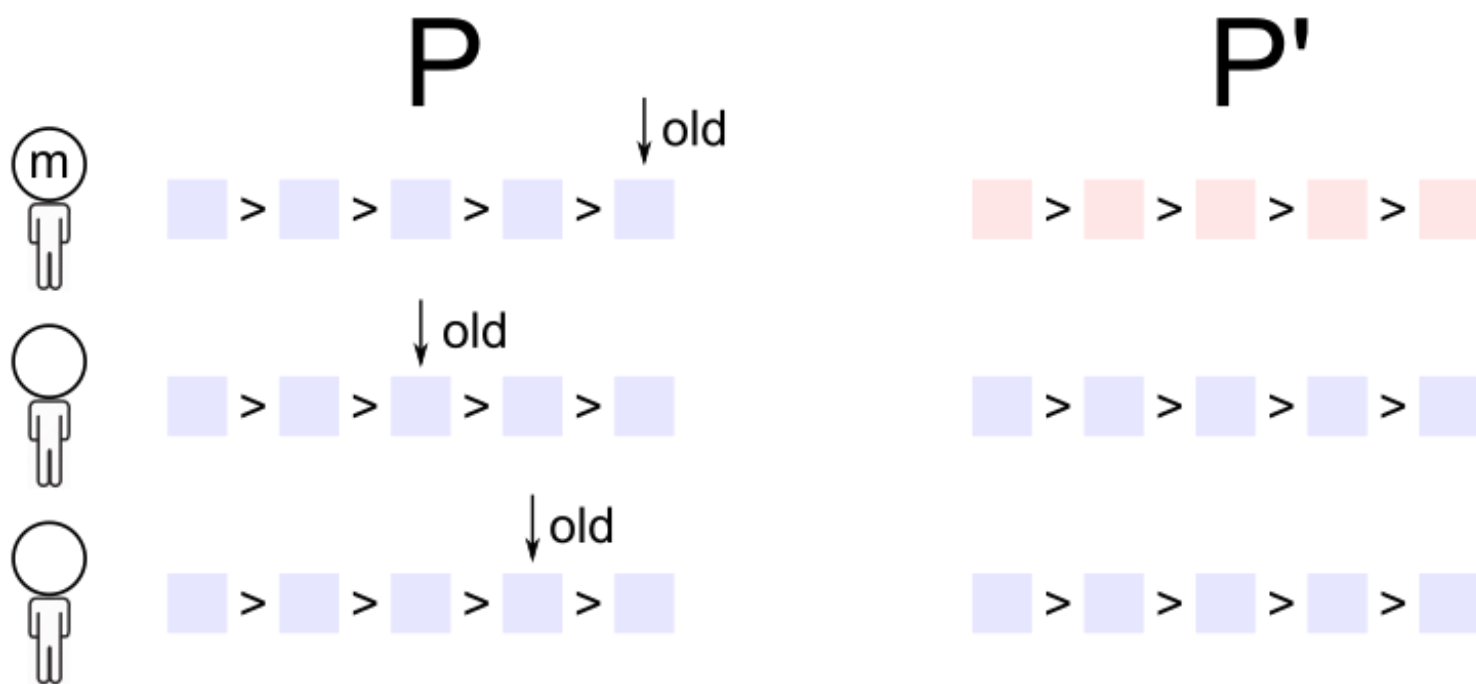
If a man proposes to a woman during DA on the profile  $P'$ , then he must also propose to her during DA on the profile  $P$ .



# 3

## No new proposal

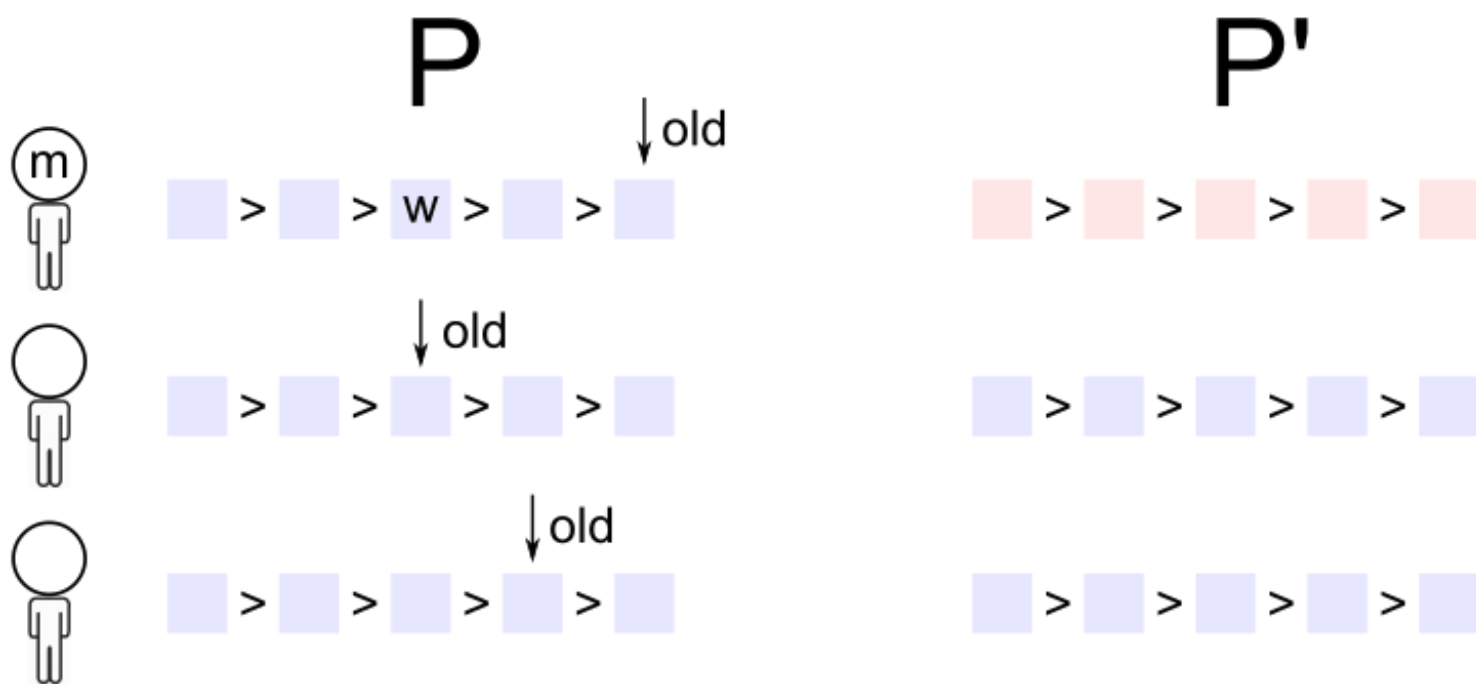
If a man proposes to a woman during DA on the profile  $P'$ , then he must also propose to her during DA on the profile  $P$ .



# 3

## No new proposal

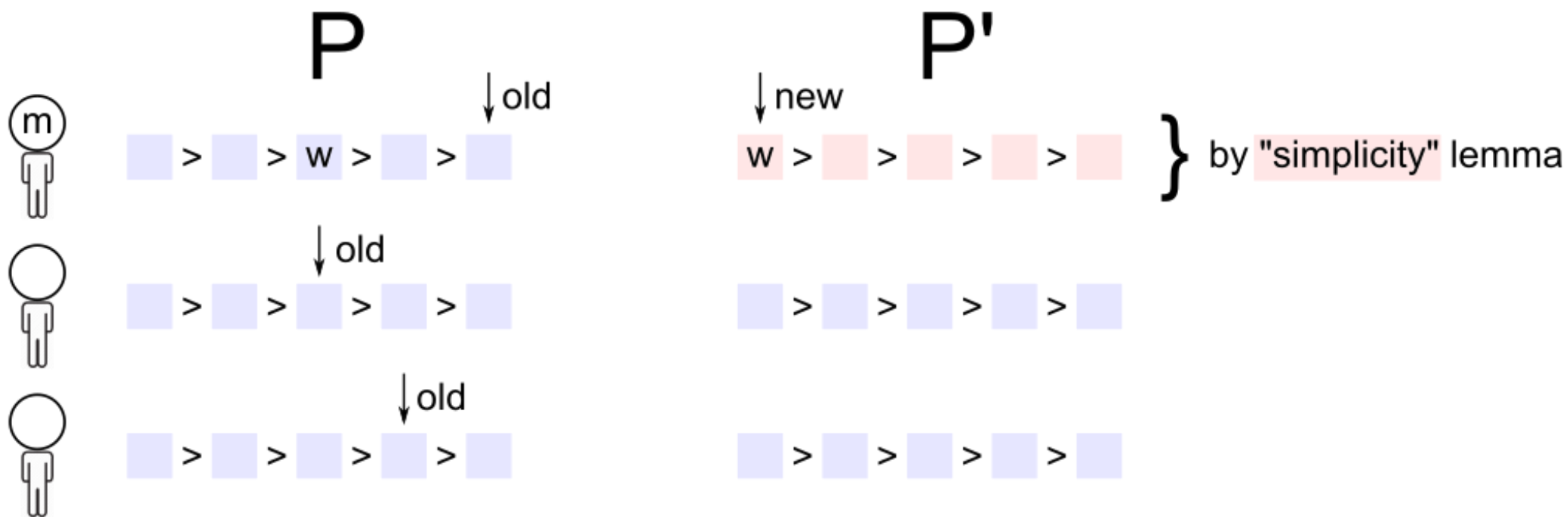
If a man proposes to a woman during DA on the profile  $P'$ , then he must also propose to her during DA on the profile  $P$ .



# 3

## No new proposal

If a man proposes to a woman during DA on the profile  $P'$ , then he must also propose to her during DA on the profile  $P$ .

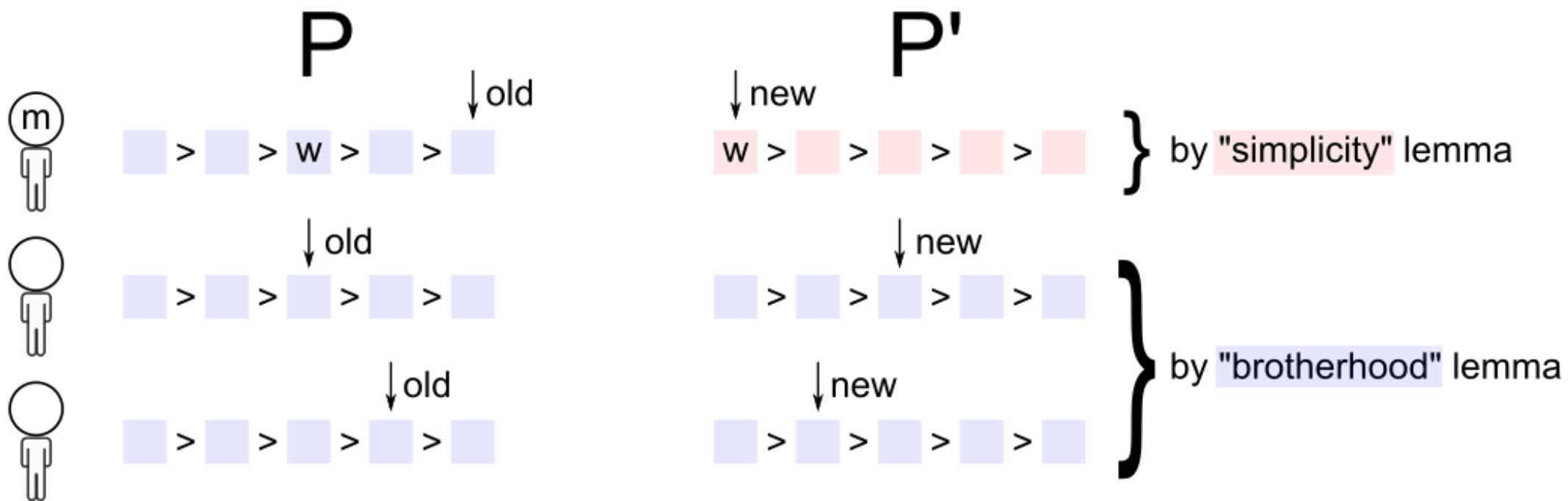




# 3

## No new proposal

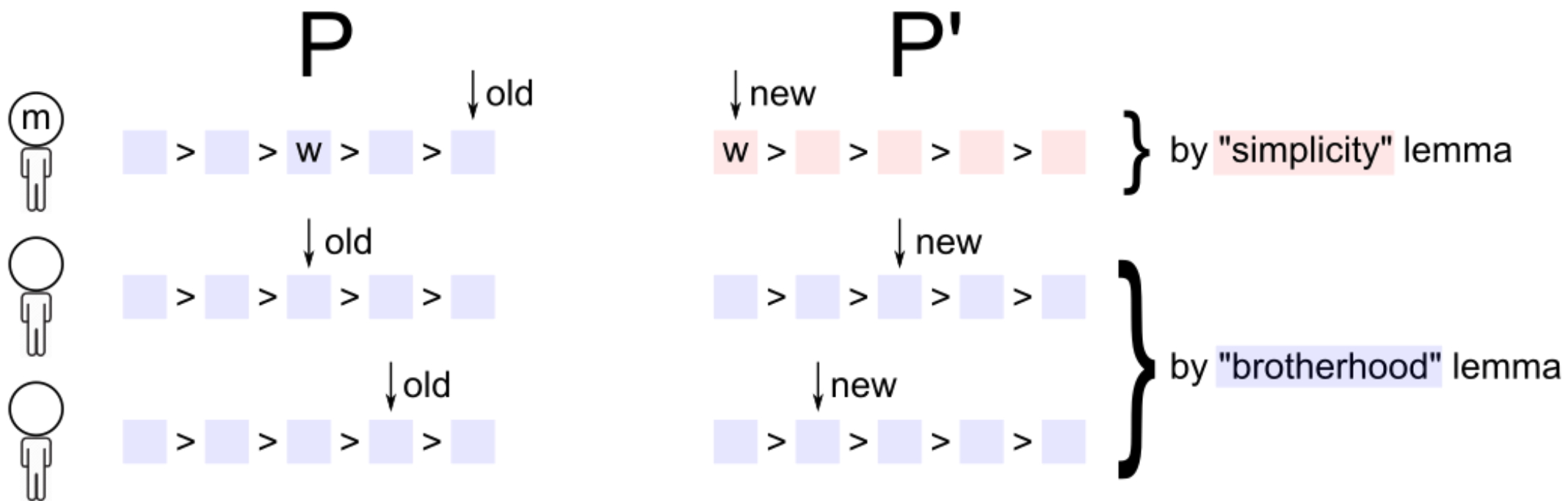
If a man proposes to a woman during DA on the profile  $P'$ , then he must also propose to her during DA on the profile  $P$ .



# 3

## No new proposal

If a man proposes to a woman during DA on the profile  $P'$ , then he must also propose to her during DA on the profile  $P$ .



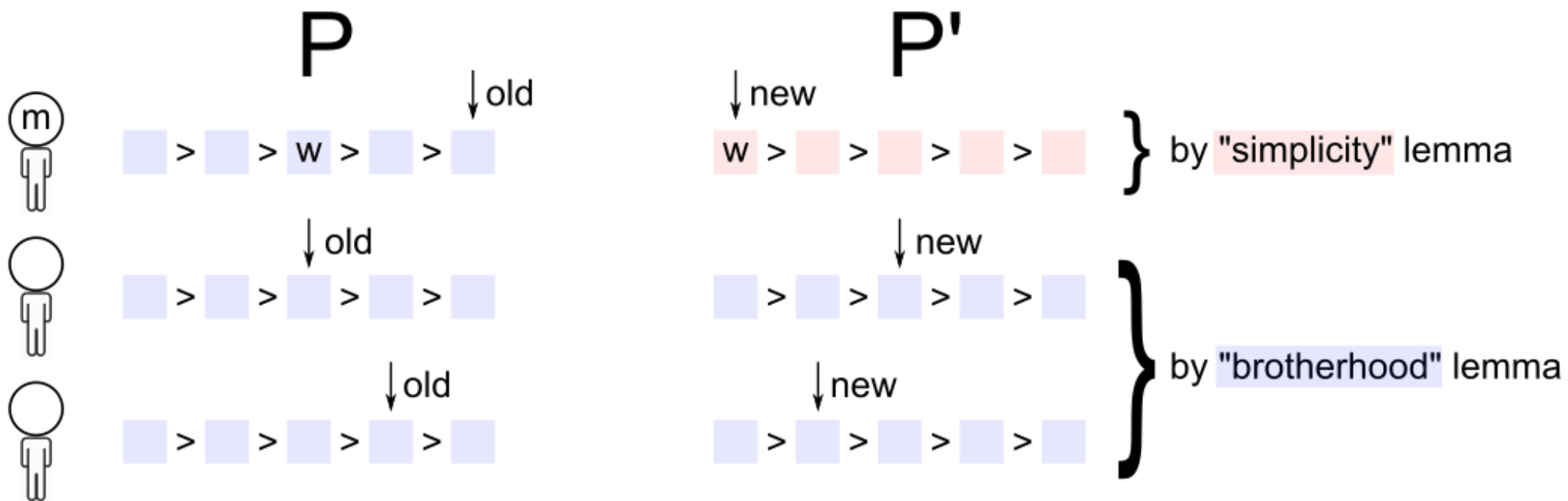
As a consequence:

If a woman receives just one proposal during DA on  $P$ , then she receives only one proposal (from the same man) during DA on  $P'$ .

# 3

## No new proposal

If a man proposes to a woman during DA on the profile  $P'$ , then he must also propose to her during DA on the profile  $P$ .



As a consequence:

If a woman receives just one proposal during DA on  $P$ , then she receives only one proposal (from the same man) during DA on  $P'$ .



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

DA algorithm is strategyproof for the men.

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

Consider the execution of DA on the true profile  $P$ . Recall that  $X = DA(P)$ .

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

Consider the execution of DA on the true profile  $P$ . Recall that  $X = DA(P)$ .

Suppose manipulator  $m$  proposes to his  $X$ -partner (i.e.,  $X(m)$ ) in round  $k$  of the algorithm.

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

Consider the execution of DA on the true profile  $P$ . Recall that  $X = DA(P)$ .

Suppose manipulator  $m$  proposes to his  $X$ -partner (i.e.,  $X(m)$ ) in round  $k$  of the algorithm.

We will show that:

Any man who proposes to his  $X$ -partner in round  $k$  or later is matched to his  $X$ -partner under  $X'$  as well.



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

Consider the execution of DA on the true profile  $P$ . Recall that  $X = DA(P)$ .

Suppose manipulator  $m$  proposes to his  $X$ -partner (i.e.,  $X(m)$ ) in round  $k$  of the algorithm.

We will show that:

Any man who proposes to his  $X$ -partner in round  $k$  or later is matched to his  $X$ -partner under  $X'$  as well.

Then,  $X(m) = X'(m)$ , which means the manipulator does not improve.

Any man who proposes to his  $X$ -partner in round  $k$  or later is matched to his  $X$ -partner under  $X'$  as well.

Any man who proposes to his  $X$ -partner in round  $k$  or later is matched to his  $X$ -partner under  $X'$  as well.

Proof by induction.

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Base case:

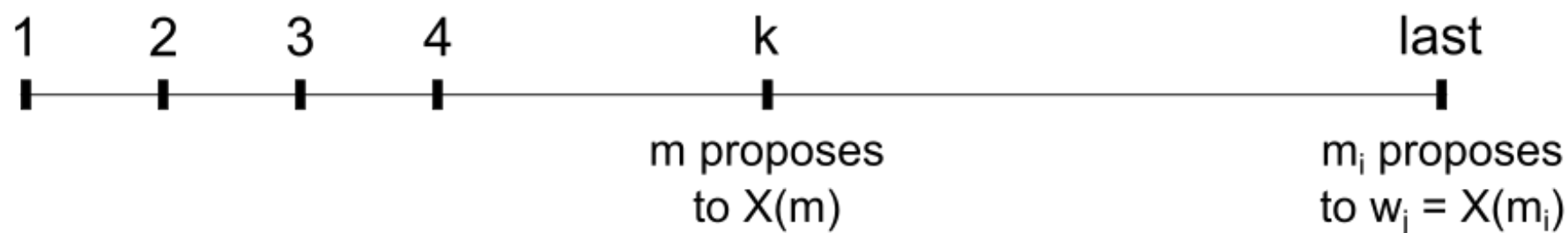
Consider the man  $m_i$  who proposes to his X-partner  $w_j$  in the **last** round of DA on profile  $P$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Base case:

Consider the man  $m_i$  who proposes to his X-partner  $w_j$  in the **last** round of DA on profile  $P$ .

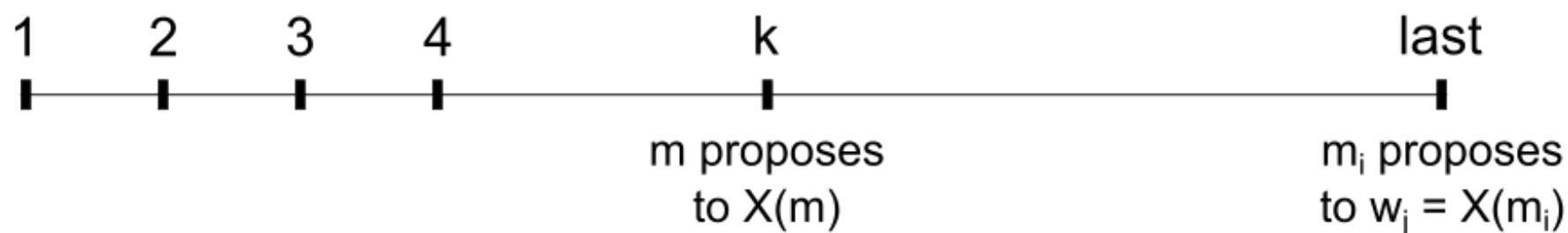


Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Base case:

Consider the man  $m_i$  who proposes to his X-partner  $w_j$  in the **last** round of DA on profile  $P$ .



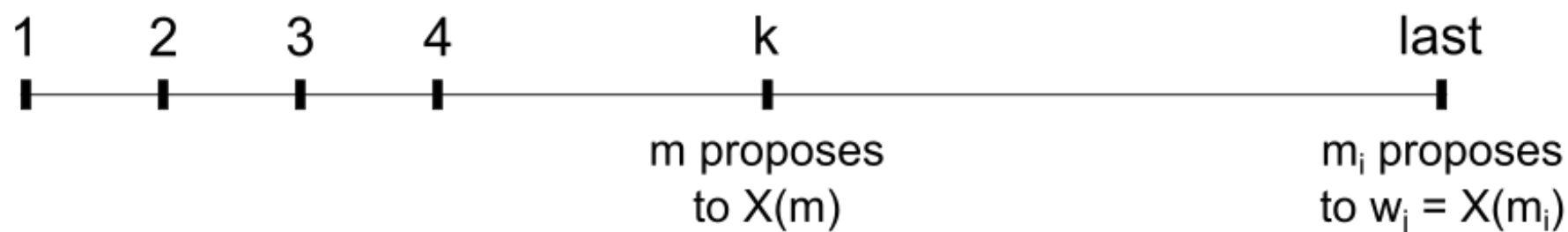
Then,  $m_i$  is the *only* proposal  $w_j$  receives during DA( $P$ )

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Base case:

Consider the man  $m_i$  who proposes to his X-partner  $w_j$  in the **last** round of DA on profile  $P$ .



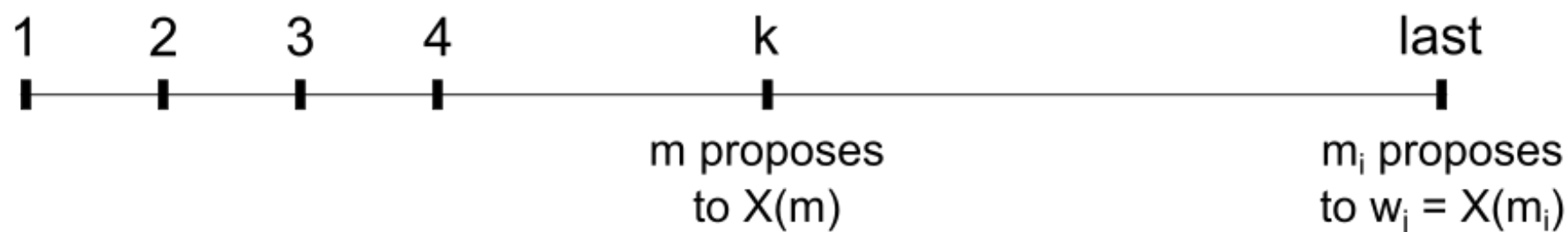
Then,  $m_i$  is the *only* proposal  $w_j$  receives during DA( $P$ )  
(since otherwise the man rejected by  $w_j$  will propose to his X-partner in a later round).

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Base case:

Consider the man  $m_i$  who proposes to his X-partner  $w_j$  in the **last** round of DA on profile  $P$ .



Then,  $m_i$  is the *only* proposal  $w_j$  receives during DA( $P$ )  
(since otherwise the man rejected by  $w_j$  will propose to his X-partner in a later round).

By "no new proposal" lemma,  $w_j$  receives only one proposal during DA( $P'$ ).

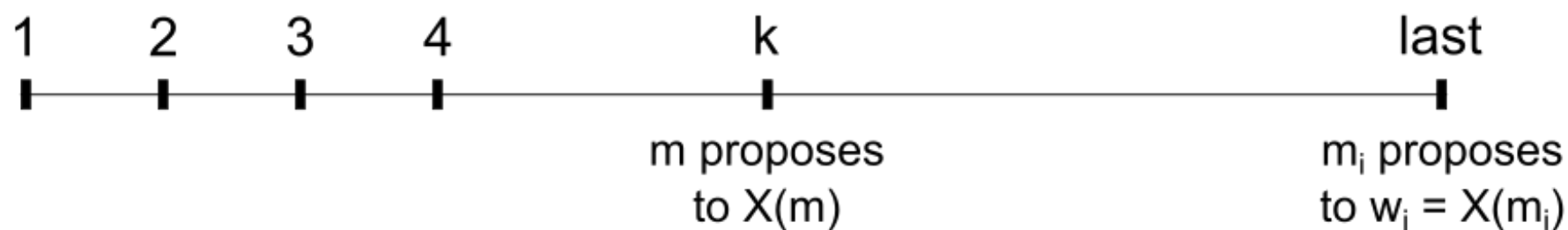


Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Base case:

Consider the man  $m_i$  who proposes to his X-partner  $w_j$  in the **last** round of DA on profile  $P$ .



Then,  $m_i$  is the *only* proposal  $w_j$  receives during DA( $P$ )  
(since otherwise the man rejected by  $w_j$  will propose to his X-partner in a later round).

By "no new proposal" lemma,  $w_j$  receives only one proposal during DA( $P'$ ).

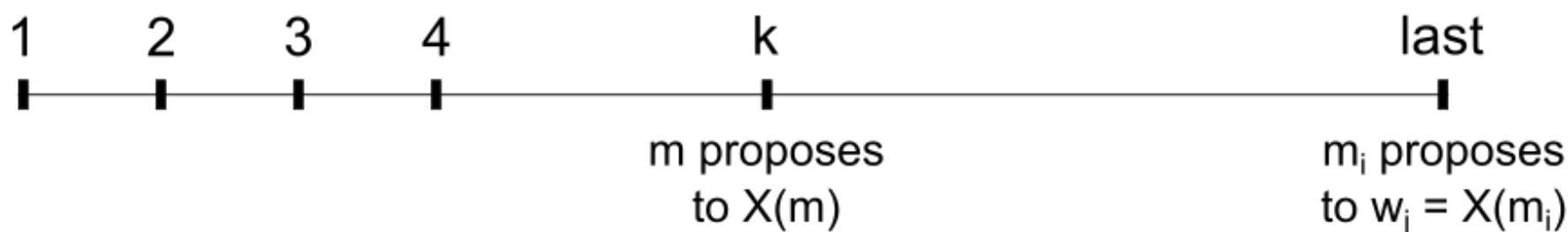
Then,  $w_j = X'(m_i)$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Base case:

Consider the man  $m_i$  who proposes to his X-partner  $w_j$  in the **last** round of DA on profile  $P$ .



Then,  $m_i$  is the *only* proposal  $w_j$  receives during DA( $P$ )  
(since otherwise the man rejected by  $w_j$  will propose to his X-partner in a later round).

By "no new proposal" lemma,  $w_j$  receives only one proposal during DA( $P'$ ).

Then,  $w_j = X'(m_i)$ .



Any man who proposes to his  $X$ -partner in round  $k$  or later is matched to his  $X$ -partner under  $X'$  as well.

Proof by induction.

Any man who proposes to his  $X$ -partner in round  $k$  or later is matched to his  $X$ -partner under  $X'$  as well.

Proof by induction.

Induction step:

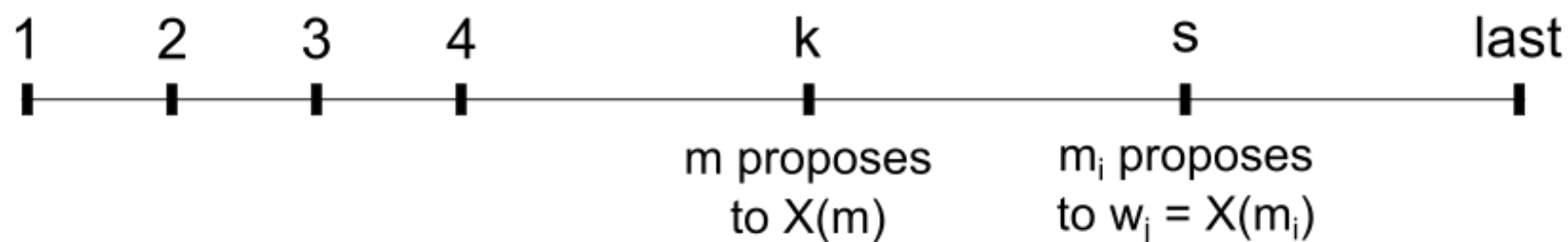
Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .

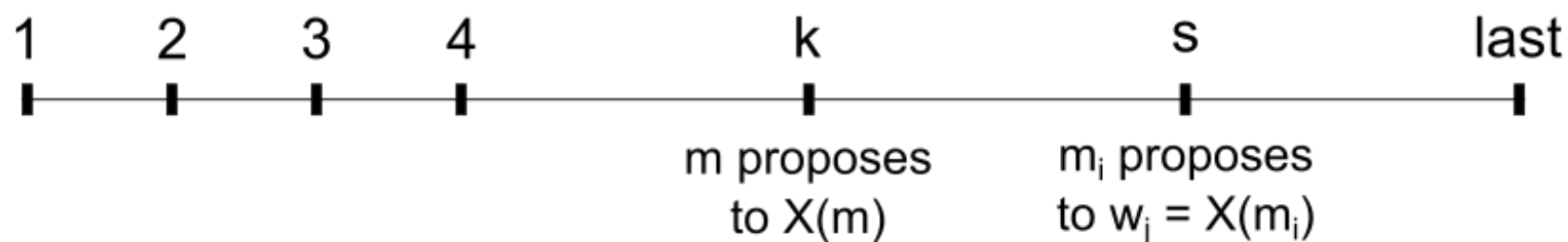


Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



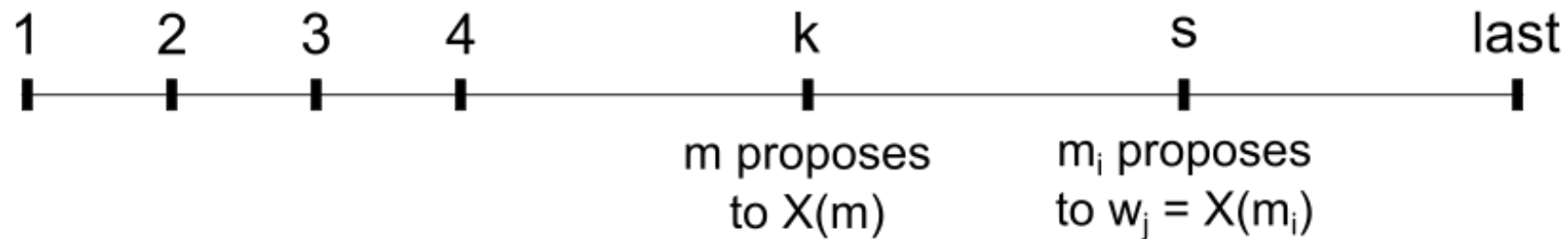
Suppose man  $m_i$  proposes to his X-partner  $w_j = X(m_i)$  in round  $s$  of DA on profile  $P$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Suppose man  $m_i$  proposes to his X-partner  $w_j = X(m_i)$  in round  $s$  of DA on profile  $P$ .

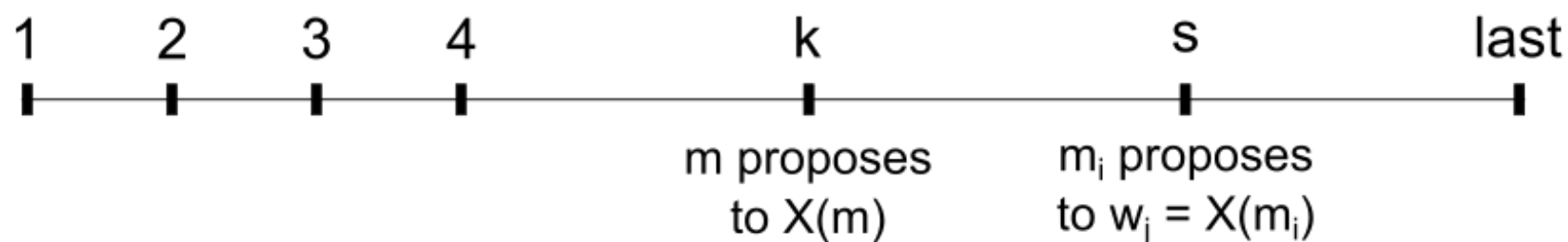
We want to show that  $w_j = X'(m_i)$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



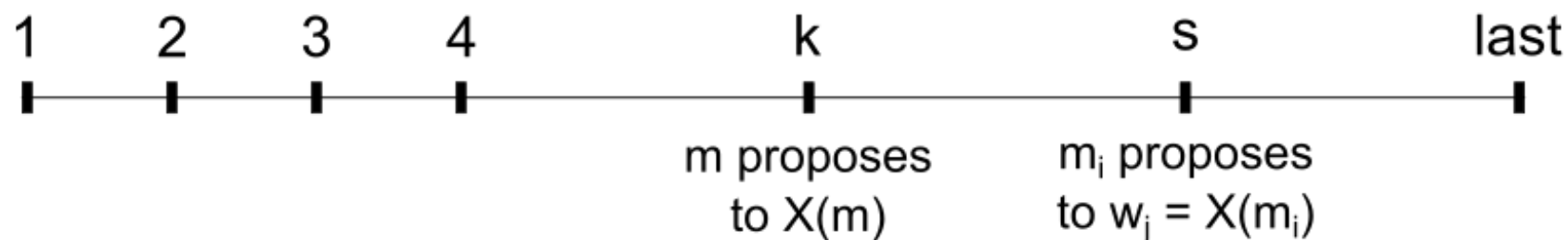


Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



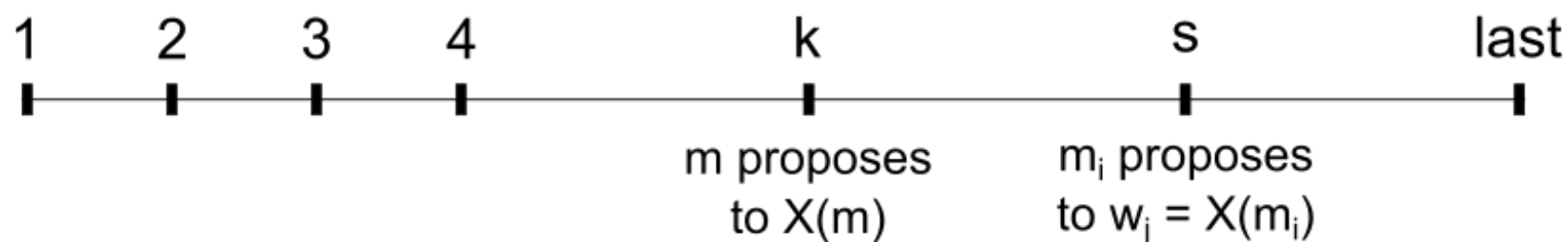
Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

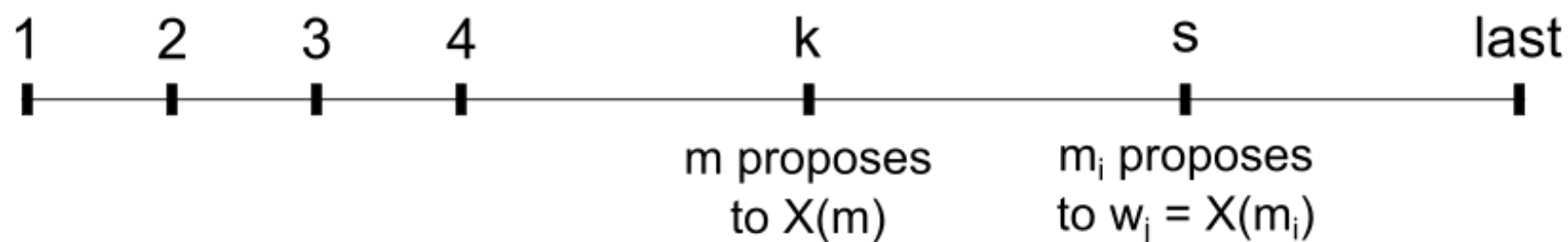
If  $R$  is empty, then  $m_i$  is the only proposal that  $w_j$  receives.

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

If  $R$  is empty, then  $m_i$  is the only proposal that  $w_j$  receives.

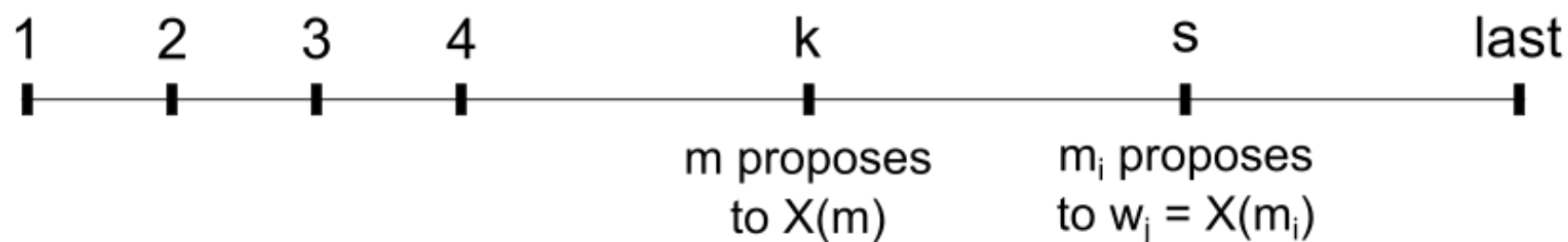
By "no new proposal" lemma,  $w_j = X'(m_i)$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



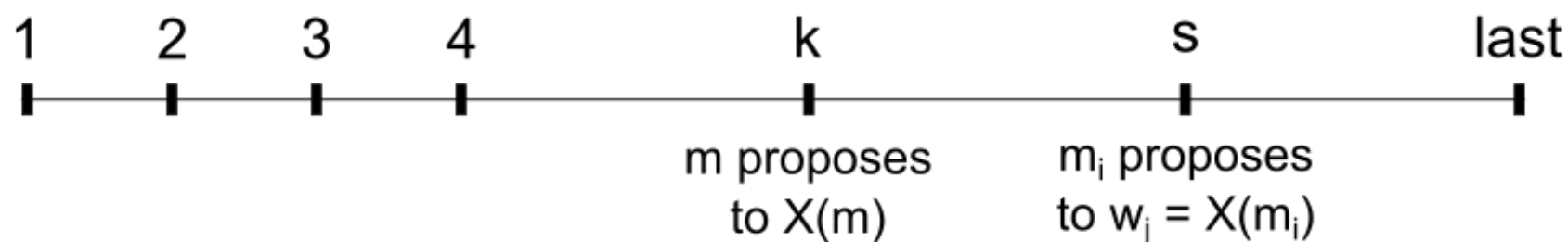
Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

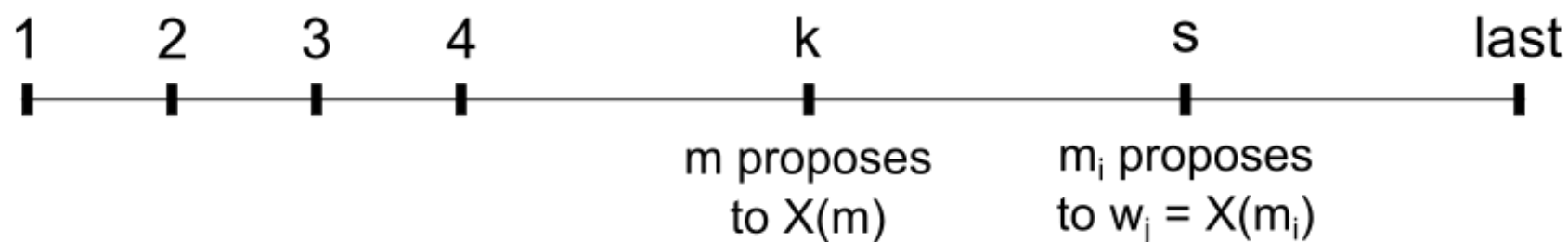
If  $R$  is non-empty, then let  $m_F$  be  $w_j$ 's favorite man in  $R$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

If  $R$  is non-empty, then let  $m_F$  be  $w_j$ 's favorite man in  $R$ .

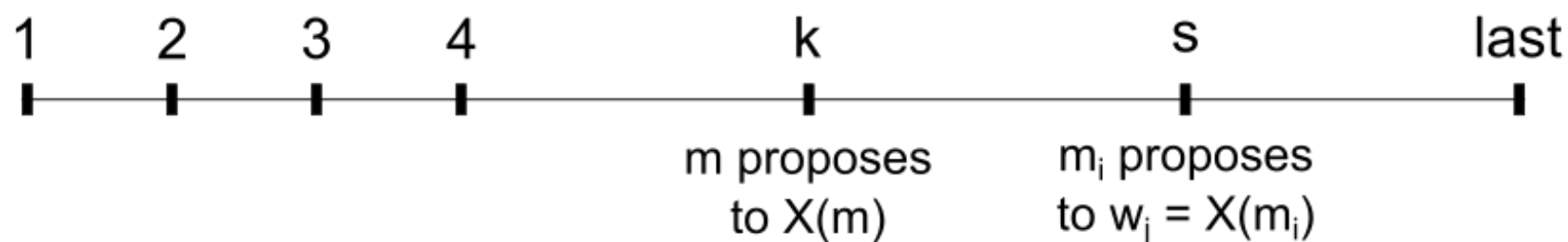
Then,  $m_F$  proposes to his X-partner in round  $s+1$  or later. Thus,  $X(m_F) = X'(m_F)$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

If  $R$  is non-empty, then let  $m_F$  be  $w_j$ 's favorite man in  $R$ .

Then,  $m_F$  proposes to his X-partner in round  $s+1$  or later. Thus,  $X(m_F) = X'(m_F)$ .

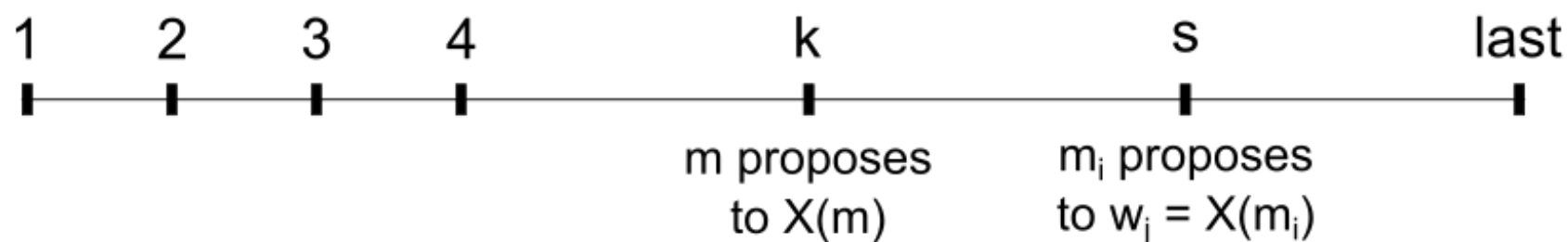
This also means that  $m_F$  can't be the manipulator  $m$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



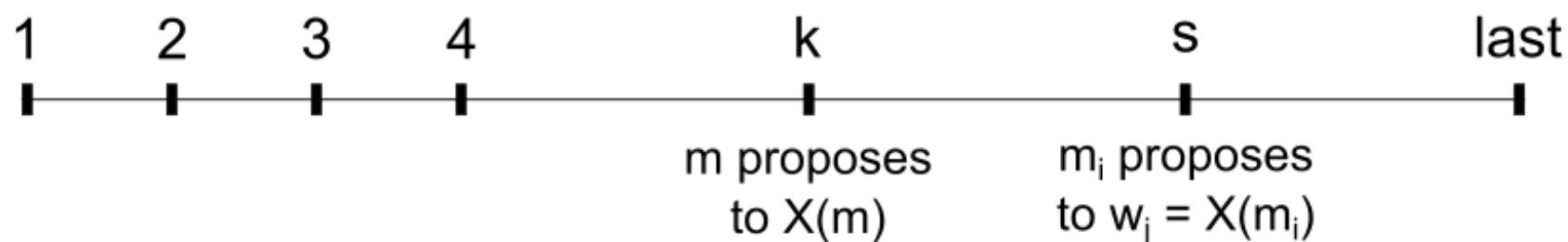


Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

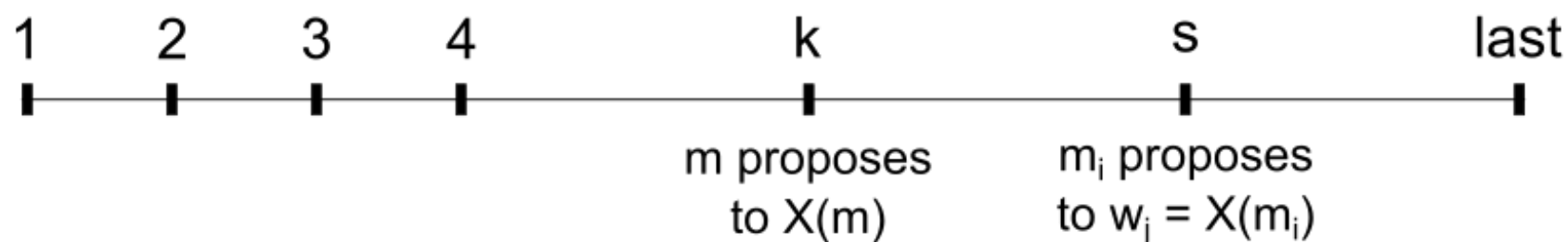
Since  $m_F$  is truthful, his preferences are unchanged between the profiles  $P$  and  $P'$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

Since  $m_F$  is truthful, his preferences are unchanged between the profiles  $P$  and  $P'$ .

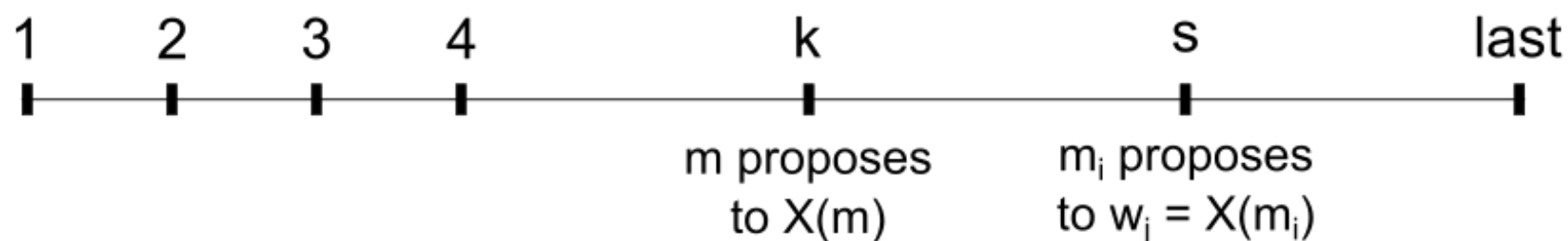
So,  $m_F$  must propose to  $w_j$  during DA on  $P'$ , and is again rejected since  $X(m_F) = X'(m_F)$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Let  $R$  = set of men rejected by  $w_j$  during DA on profile  $P$  (across ALL rounds).

Since  $m_F$  is truthful, his preferences are unchanged between the profiles  $P$  and  $P'$ .

So,  $m_F$  must propose to  $w_j$  during DA on  $P'$ , and is again rejected since  $X(m_F) = X'(m_F)$ .

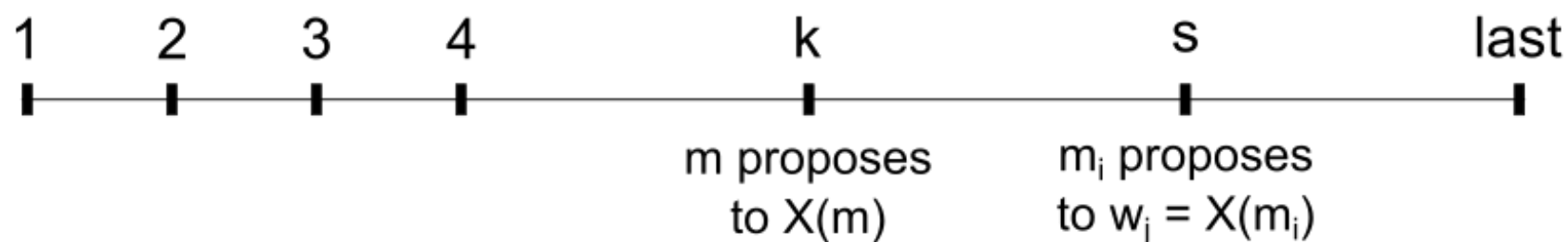
Thus,  $w_j$  receives at least one more proposal (besides  $m_F$ ) during DA( $P'$ ).

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .

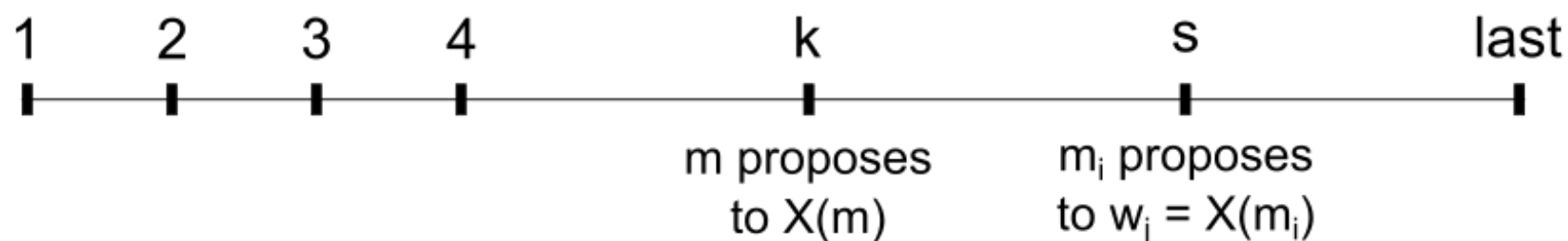


Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



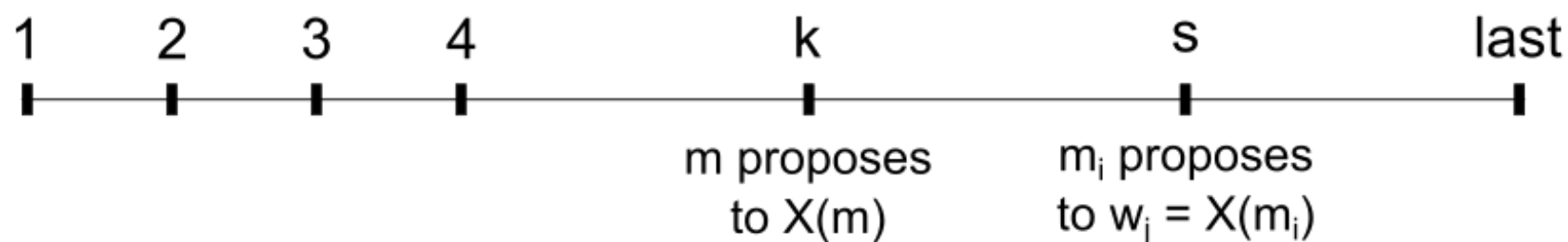
By "no new proposal" lemma, the man  $m' = X'(w_j)$  proposes to  $w_j$  during  $DA(P)$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



By "no new proposal" lemma, the man  $m' = X'(w_j)$  proposes to  $w_j$  during  $DA(P)$ .

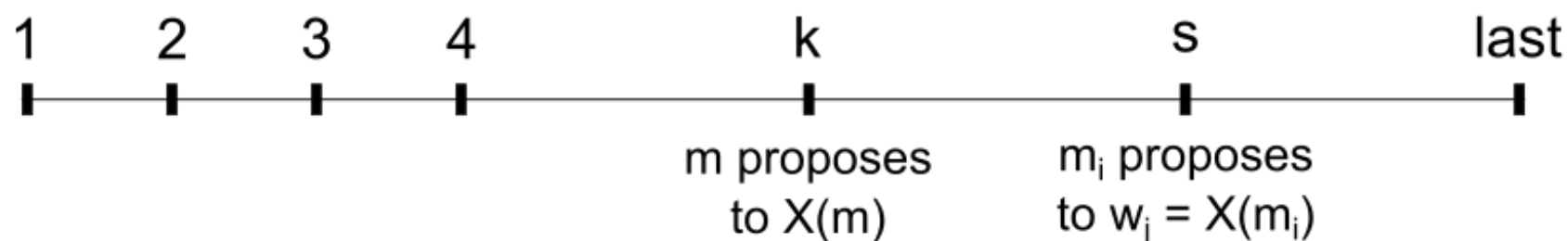
- If  $m' = X(w_j) = m_i$ , then we are done.

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



By "no new proposal" lemma, the man  $m' = X'(w_j)$  proposes to  $w_j$  during  $DA(P)$ .

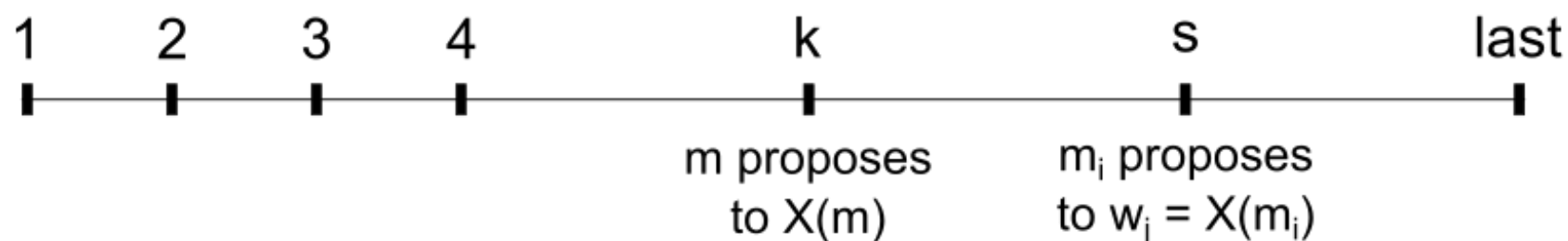
- If  $m' = X(w_j) = m_i$ , then we are done.
- If  $m' >_w m_i$ , then  $w_j$  would have rejected her X-partner during  $DA(P)$ ---a contradiction.

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



By "no new proposal" lemma, the man  $m' = X'(w_j)$  proposes to  $w_j$  during  $DA(P)$ .

- If  $m' = X(w_j) = m_i$ , then we are done.
- If  $m' >_w m_i$ , then  $w_j$  would have rejected her X-partner during  $DA(P)$ ---a contradiction.
- If  $m_i >_w m'$ , then  $w_j$  weakly prefers  $m_i$  over  $m'$ , and would have rejected  $m'$  during  $DA(P')$ , again a contradiction.

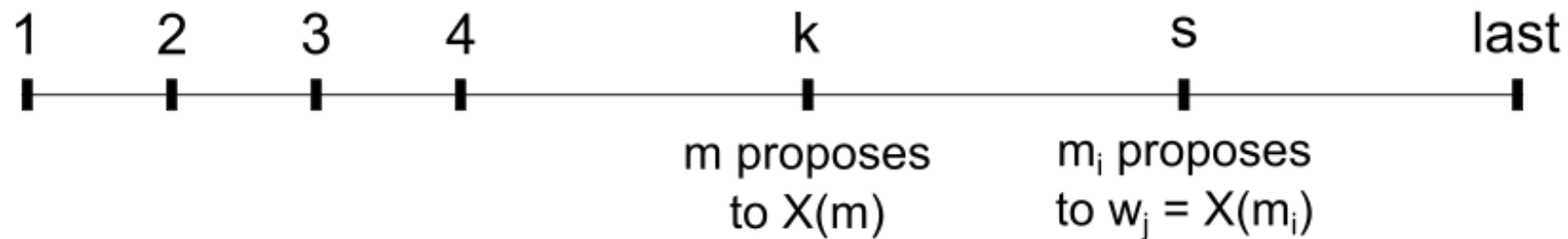


Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .

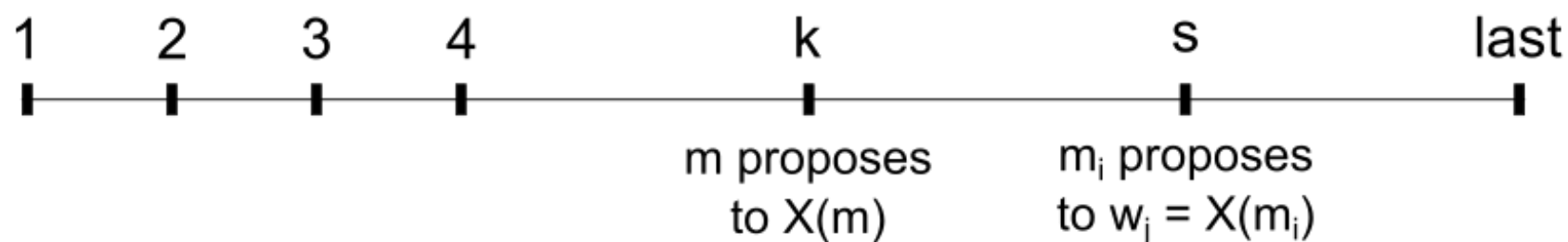


Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



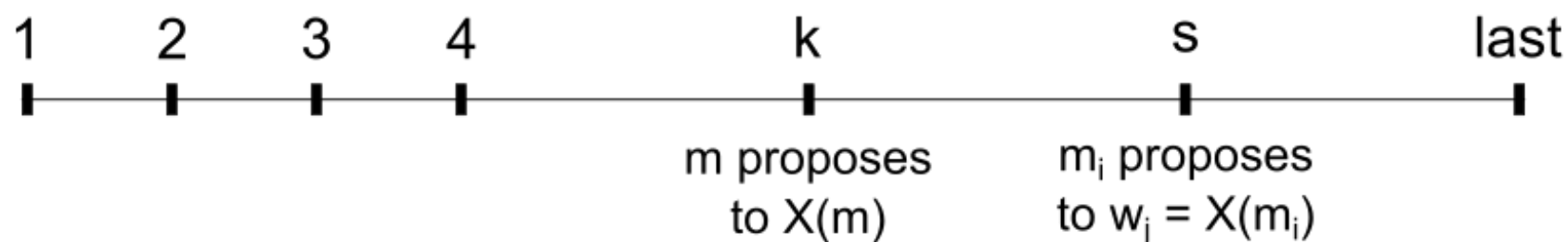
Therefore,  $X(m_j) = w_j = X'(m_j)$ .

Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Therefore,  $X(m_j) = w_j = X'(m_j)$ .

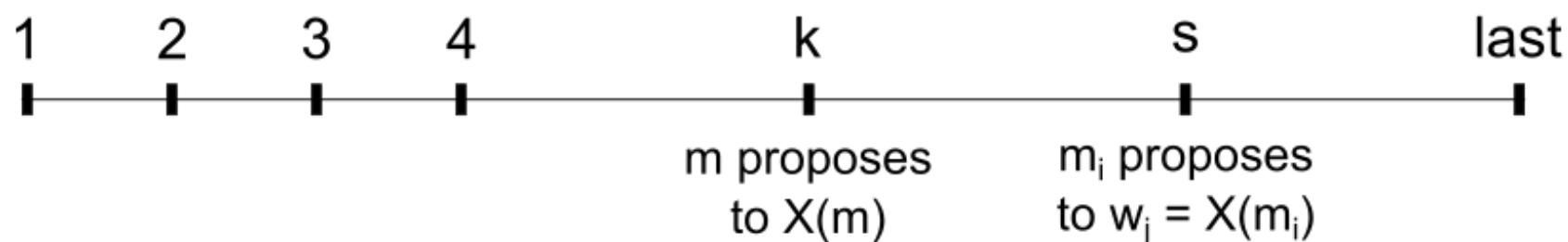


Any man who proposes to his X-partner in round  $k$  or later is matched to his X-partner under  $X'$  as well.

Proof by induction.

Induction step:

Suppose the claim holds for all rounds  $s+1$  or later, and we want to prove it for round  $s$ .



Therefore,  $X(m_j) = w_j = X'(m_j)$ .



[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

Consider the execution of DA on the true profile  $P$ . Recall that  $X = DA(P)$ .

Suppose man  $m$  proposes to his  $X$ -partner (i.e.,  $X(m)$ ) in round  $k$  of the algorithm.

Any man who proposes to his  $X$ -partner in round  $k$  or later is matched to his  $X$ -partner under  $X'$  as well.

Then,  $X(m) = X'(m)$ , which means the manipulator does not improve.

[Dubins and Freedman, *Amer. Math. Mon.* 1981; Roth, *MOR* 1982]

**DA algorithm is strategyproof for the men.**

Suppose, for contradiction, that DA can be manipulated by a man  $m$  on the profile  $P$ .

Consider the execution of DA on the true profile  $P$ . Recall that  $X = DA(P)$ .

Suppose man  $m$  proposes to his  $X$ -partner (i.e.,  $X(m)$ ) in round  $k$  of the algorithm.

Any man who proposes to his  $X$ -partner in round  $k$  or later is matched to his  $X$ -partner under  $X'$  as well.

Then,  $X(m) = X'(m)$ , which means the manipulator does not improve.



