Sisterhood in the Gale-Shapley Matching Algorithm

Yannai A. Gonczarowski

Einstein Institute of Mathematics and Center for the Study of Rationality
The Hebrew University of Jerusalem

June 3, 2013

Joint work with Ehud Friedgut

The Electronic Journal of Combinatorics **20(2)** (2013), #P12

The Stable Matching Problem

- Two disjoint finite sets to be matched: women W and men M.
  - Assume 1-to-1 for now.
  - Assume $|W| = |M|$ for now.

The Stable Matching Problem

- Two disjoint finite sets to be matched: women W and men M.
  - Assume 1-to-1 for now.
  - Assume $|W| = |M|$ for now.
- Preferences for each woman and for each man.
  - Assume a strict order of preference for each woman over all men and vice versa.
  - Assume no blacklists for now.

The Stable Matching Problem

- Two disjoint finite sets to be matched: women W and men M.
  - Assume 1-to-1 for now.
  - Assume $|W| = |M|$ for now.
- Preferences for each woman and for each man.
  - Assume a strict order of preference for each woman over all men and vice versa.
  - Assume no blacklists for now.
- The goal: a stable matching.
  - If w and m are matched, and if w' and m' are matched, then w and m' should not both prefer each other over their spouses.

The Gale-Shapley Deferred-Acceptance Algorithm

Gale and Shapley (1962)

The following algorithm yields a stable matching.

1. On each night, every man serenades under the window of the woman he prefers most out of those who have not yet rejected him.

The Gale-Shapley Deferred-Acceptance Algorithm

Gale and Shapley (1962)

The following algorithm yields a stable matching.

1. On each night, every man serenades under the window of the woman he prefers most out of those who have not yet rejected him.
2. On each night, every woman rejects all the men serenading under her window, except for the one she prefers most among them.

The Gale-Shapley Deferred-Acceptance Algorithm

Gale and Shapley (1962)

The following algorithm yields a stable matching.

1. On each night, every man serenades under the window of the woman he prefers most out of those who have not yet rejected him.
2. On each night, every woman rejects all the men serenading under her window, except for the one she prefers most among them.
3. When no more rejections occur, each woman is matched with the man serenading under her window.

Gender Duality and Manipulation Incentives

Gale and Shapley (1962)

No stable matching is better for any man.

McVitie and Wilson (1971)

No stable matching is worse for any woman.

Gender Duality and Manipulation Incentives

Gale and Shapley (1962)

No stable matching is better for any man.

McVitie and Wilson (1971)

No stable matching is worse for any woman.

Dubins and Freedman (1981)

No subset of men can lie in a way that would make them all better off lying.

Gale and Sotomayor (1985)

Generally, there exists a woman who would be better off lying.

Gender Duality and Manipulation Incentives

Gale and Shapley (1962)

No stable matching is better for any man.

McVitie and Wilson (1971)

No stable matching is worse for any woman.

Dubins and Freedman (1981)

No subset of men can lie in a way that would make them all better off lying.

Gale and Sotomayor (1985)

Generally, there exists a woman who would be better off lying.

Note: the latter two do not follow from the former two.

Example: Manipulation by Women

Men's Preferences

m_1	w_2	w_1	w_3
m_2	w_1	w_2	w_3
m_3	w_1	w_3	w_2

Women's Preferences

w_1	$m_1 > m_2 > m_3$
w_2	$m_2 > m_1$
w_3	any

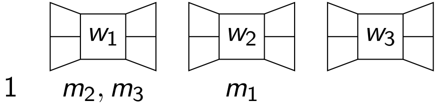
Example: Manipulation by Women

Men's Preferences

m_1	w_2	w_1	w_3
m_2	w_1	w_2	w_3
m_3	w_1	w_3	w_2

Women's Preferences

w_1	$m_1 > m_2 > m_3$
w_2	$m_2 > m_1$
w_3	any



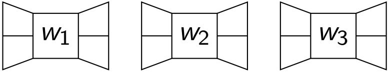
Example: Manipulation by Women

Men's Preferences

m_1	w_2	w_1	w_3
m_2	w_1	w_2	w_3
m_3	w_1	w_3	w_2

Women's Preferences

w_1	$m_1 > m_2 > m_3$
w_2	$m_2 > m_1$
w_3	any



	w_1	w_2	w_3
1	m_2, m_3	m_1	
2	m_2	m_1	m_3

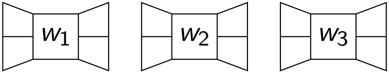
Example: Manipulation by Women

Men's Preferences

m_1	w_2	w_1	w_3
m_2	w_1	w_2	w_3
m_3	w_1	w_3	w_2

Women's Preferences

w_1	$m_1 > m_{\not{2}3} > m_{\not{3}2}$
w_2	$m_2 > m_1$
w_3	any



	w_1	w_2	w_3
1	m_2, m_3	m_1	
2	m_2	m_1	m_3

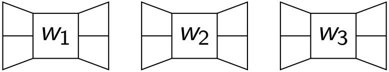
Example: Manipulation by Women

Men's Preferences

m_1	w_2	w_1	w_3
m_2	w_1	w_2	w_3
m_3	w_1	w_3	w_2

Women's Preferences

w_1	$m_1 > m_{\not{2}3} > m_{\not{3}2}$
w_2	$m_2 > m_1$
w_3	any



	w_1	w_2	w_3
1	m_2, m_3	m_1	
2	m_2	m_1	m_3
2	m_3	m_1, m_2	

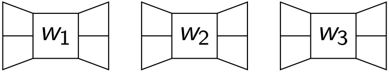
Example: Manipulation by Women

Men's Preferences

m_1	w_2	w_1	w_3
m_2	w_1	w_2	w_3
m_3	w_1	w_3	w_2

Women's Preferences

w_1	$m_1 > m_{\not{2}3} > m_{\not{3}2}$
w_2	$m_2 > m_1$
w_3	any



	w_1	w_2	w_3
1	m_2, m_3	m_1	
2	m_2	m_1	m_3
2	m_3	m_1, m_2	
3	m_1, m_3	m_2	

Example: Manipulation by Women

Men's Preferences

m_1	w_2	w_1	w_3
m_2	w_1	w_2	w_3
m_3	w_1	w_3	w_2

Women's Preferences

w_1	$m_1 > m_{\cancel{2}3} > m_{\cancel{3}2}$
w_2	$m_2 > m_1$
w_3	any

	w_1	w_2	w_3
1	m_2, m_3	m_1	
2	m_2	m_1	m_3
2	m_3	m_1, m_2	
3	m_1, m_3	m_2	
4	m_1	m_2	m_3

Example: Manipulation by Women

Men's Preferences

m_1	w_2	w_1	w_3
m_2	w_1	w_2	w_3
m_3	w_1	w_3	w_2

Women's Preferences

w_1	$m_1 > m_{\cancel{2}3} > m_{\cancel{3}2}$
w_2	$m_2 > m_1$
w_3	any

	w_1	w_2	w_3
1	m_2, m_3	m_1	
2	m_2	m_1	m_3
2	m_3	m_1, m_2	
3	m_1, m_3	m_2	
4	m_1	m_2	m_3

- w_1 improved her match, but so did w_2; and w_3 is unharmed.

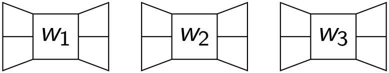
Example: Manipulation by Women

Men's Preferences

m_1	w_2	w_1	w_3
m_2	w_1	w_2	w_3
m_3	w_1	w_3	w_2

Women's Preferences

w_1	$m_1 > m_{\not{2}3} > m_{\not{3}2}$
w_2	$m_2 > m_1$
w_3	any



	w_1	w_2	w_3
1	m_2, m_3	m_1	
2	m_2	m_1	m_3
2	m_3	m_1, m_2	
3	m_1, m_3	m_2	
4	m_1	m_2	m_3

- w_1 improved her match, but so did w_2; and w_3 is unharmed.
- w_1 made w_2 "give up" m_1 by making sure w_2 is approached by someone w_2 prefers better.

Sisterhood Theorem

Assume that a subset of the women declare false orders of preference for themselves.

We examines two runs of the Gale-Shapley algorithm:

- OA — according to everyone's true preferences; yields the matching O.
- NA — according to the liars' false preferences, and everyone else's true preferences; yields the matching N.

Sisterhood Theorem

Assume that a subset of the women declare false orders of preference for themselves.

We examines two runs of the Gale-Shapley algorithm:

- OA — according to everyone's true preferences; yields the matching O.
- NA — according to the liars' false preferences, and everyone else's true preferences; yields the matching N.

Theorem (Sisterhood)

Under the above conditions, if all lying women are weakly better off, then:

1. *All women are weakly better off.*
2. *All men are weakly worse off.*

Sisterhood Theorem

Assume that a subset of the women declare false orders of preference for themselves.

We examines two runs of the Gale-Shapley algorithm:

- OA — according to everyone's true preferences; yields the matching O.
- NA — according to the liars' false preferences, and everyone else's true preferences; yields the matching N.

Theorem (Sisterhood)

Under the above conditions, if all lying women are weakly better off, then:

1. *All women are weakly better off.*
2. *All men are weakly worse off.*

No such "hoodness" exists within any other subset of $W \cup M$. Indeed, when even a single man lies and is weakly b/o, some women and men may be b/o, and some others — w/o.

An Easy Proof?

Observation

If every lying woman w lies in an optimal way (i.e. the lies constitute a Nash Equilibrium in the lying game), then the new matching is stable.

Background

Monogamous Case

Polygamy, Blacklists, and Mismatched Quotas

An Easy Proof?

Observation

If every lying woman w lies in an optimal way (i.e. the lies constitute a Nash Equilibrium in the lying game), then the new matching is stable.

Proof.

The new matching is obviously stable w.r.t. the new preferences. It is thus enough to consider couples in which at least one liar participates.

$$\begin{array}{cc} w & w' \\ | & | \\ m & m' \end{array}$$

□

An Easy Proof?

Observation

If every lying woman w lies in an optimal way (i.e. the lies constitute a Nash Equilibrium in the lying game), then the new matching is stable.

Proof.

The new matching is obviously stable w.r.t. the new preferences. It is thus enough to consider couples in which at least one liar participates.

$$\begin{array}{cc} w & w' \\ | & | \\ m & m' \end{array}$$

□

So, what's the problem? Why would someone lie in a non-optimal way? Why do we care about non-equilibrium?

When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_1 > m_2, m_4$
w_2	$m_3 > m_1 > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any

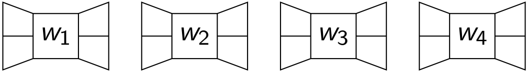
When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_1 > m_2, m_4$
w_2	$m_3 > m_1 > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any



1 m_4, m_1 m_2 m_3

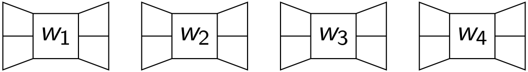
When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_1 > m_2, m_4$
w_2	$m_3 > m_1 > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any



	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4

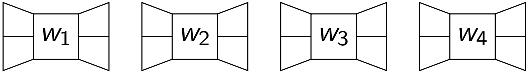
When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_{\not{1}4} > m_2, m_{\not{4}1}$
w_2	$m_3 > m_1 > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any



	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4

When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_{\not{1}4} > m_2, m_{\not{4}1}$
w_2	$m_3 > m_1 > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any

	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4
2	m_4	m_2	m_1, m_3	

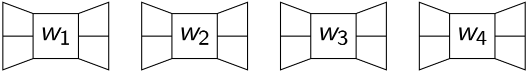
When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_{\not{1}4} > m_2, m_{\not{4}1}$
w_2	$m_3 > m_1 > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any



	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4
2	m_4	m_2	m_1, m_3	
3	m_4	m_2, m_3	m_1	

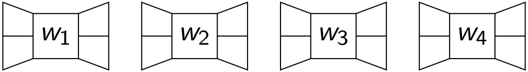
When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_{\not{1}4} > m_2, m_{\not{4}1}$
w_2	$m_3 > m_1 > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any



	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4
2	m_4	m_2	m_1, m_3	
3	m_4	m_2, m_3	m_1	
4	m_4	m_3	m_1, m_2	

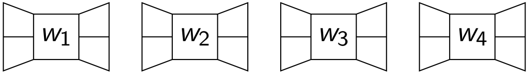
When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_{\cancel{1}4} > m_2, m_{\cancel{4}1}$
w_2	$m_3 > m_1 > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any



	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4
2	m_4	m_2	m_1, m_3	
3	m_4	m_2, m_3	m_1	
4	m_4	m_3	m_1, m_2	
5	m_4	m_1, m_3	m_2	

When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_{\not{1}4} > m_2, m_{\not{4}1}$
w_2	$m_3 > m_1 > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any

	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4
2	m_4	m_2	m_1, m_3	
3	m_4	m_2, m_3	m_1	
4	m_4	m_3	m_1, m_2	
5	m_4	m_1, m_3	m_2	
6	m_4	m_3	m_2	m_1

When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_{1 \to 4} > m_2, m_{4 \to 1}$
w_2	$m_{3 \to 1} > m_{1 \to 3} > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any

	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4
2	m_4	m_2	m_1, m_3	
3	m_4	m_2, m_3	m_1	
4	m_4	m_3	m_1, m_2	
5	m_4	m_1, m_3	m_2	
6	m_4	m_3	m_2	m_1

When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_{\not{1}4} > m_2, m_{\not{4}1}$
w_2	$m_{\not{3}1} > m_{\not{1}3} > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any

	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4
2	m_4	m_2	m_1, m_3	
3	m_4	m_2, m_3	m_1	
4	m_4	m_3	m_1, m_2	
5	m_4	m_1, m_3	m_2	
6	m_4	m_3	m_2	m_1
6	m_3, m_4	m_1	m_2	

When a Lie Need Not be Optimal

Men's Preferences

m_1	w_1	w_3	w_2	w_4
m_2	w_2	w_3	any	any
m_3	w_3	w_2	w_1	w_4
m_4	w_1	w_4	any	any

Women's Preferences

w_1	$m_3 > m_{\cancel{1}4} > m_2, m_{\cancel{4}1}$
w_2	$m_{\cancel{3}1} > m_{\cancel{1}3} > m_2, m_4$
w_3	$m_2 > m_1 > m_3$
w_4	any

	w_1	w_2	w_3	w_4
1	m_4, m_1	m_2	m_3	
2	m_1	m_2	m_3	m_4
2	m_4	m_2	m_1, m_3	
3	m_4	m_2, m_3	m_1	
4	m_4	m_3	m_1, m_2	
5	m_4	m_1, m_3	m_2	
6	m_4	m_3	m_2	m_1
6	m_3, m_4	m_1	m_2	
7	m_3	m_1	m_2	m_4

When a Lie Need Not be Optimal (cont.)

In this example:

- The truth is an optimal strategy for any coalition not including w_1.

When a Lie Need Not be Optimal (cont.)

In this example:

- The truth is an optimal strategy for any coalition not including w_1.
- No strategy for w_1 is better than the truth if all other women respond optimally to it.

When a Lie Need Not be Optimal (cont.)

In this example:

- The truth is an optimal strategy for any coalition not including w_1.
- No strategy for w_1 is better than the truth if all other women respond optimally to it.
- Thus, in no Nash equilibrium is any woman better-matched than according to all the true preferences.

When a Lie Need Not be Optimal (cont.)

In this example:

- The truth is an optimal strategy for any coalition not including w_1.
- No strategy for w_1 is better than the truth if all other women respond optimally to it.
- Thus, in no Nash equilibrium is any woman better-matched than according to all the true preferences.
- There exists a strategy for w_1 and w_2 that is better for both than the truth, but which is out-of-equilibrium due to w_2 lying suboptimally.

When a Lie Need Not be Optimal (cont.)

In this example:

- The truth is an optimal strategy for any coalition not including w_1.
- No strategy for w_1 is better than the truth if all other women respond optimally to it.
- Thus, in no Nash equilibrium is any woman better-matched than according to all the true preferences.
- There exists a strategy for w_1 and w_2 that is better for both than the truth, but which is out-of-equilibrium due to w_2 lying suboptimally.
- w_2's lie in this strategy is not equivalent to any truncation of her preferences.

One/Many-To-Many Matchings and Blacklists

One/Many-To-Many Matchings and Blacklists

- What's better off?
- What's worse off?

One/Many-To-Many Matchings and Blacklists

- What's better off?
- What's worse off?

- We still assume total preferences over individuals.
- For a person p, denote $O(p) = (o_1^p, \ldots, o_{|O(p)|}^p)$ and $N(p) = (n_1^p, \ldots, n_{|N(p)|}^p)$. Lower index = higher on p's list.

Sisterhood Theorem - Polygamous Case

Definition (Improvement)

A woman p is said to be *weakly better off* if:

1. $N(p)$ contains no-one who is blacklisted by p.
2. $|N(p)| \geq |O(p)|$.
3. For each $1 \leq i \leq O(p)$, p weakly prefers n_i^p over o_i^p.

Sisterhood Theorem - Polygamous Case

Definition (Improvement)

A woman p is said to be *weakly better off* if:

1. $N(p)$ contains no-one who is blacklisted by p.
2. $|N(p)| \geq |O(p)|$.
3. For each $1 \leq i \leq O(p)$, p weakly prefers n_i^p over o_i^p.

Definition (Worsening)

A person p is said to have *gained only worse matches* if p prefers every member of $O(p)$ over every member of $N(p) \setminus O(p)$.

Sisterhood Theorem - Polygamous Case

Definition (Improvement)

A woman p is said to be *weakly better off* if:

1. $N(p)$ contains no-one who is blacklisted by p.
2. $|N(p)| \geq |O(p)|$. (*The theorem will imply equality here.)
3. For each $1 \leq i \leq O(p)$, p weakly prefers n_i^p over o_i^p.

Definition (Worsening)

A person p is said to have *gained only worse matches* if p prefers every member of $O(p)$ over every member of $N(p) \setminus O(p)$. (*Does not require $|N(p)| \leq |O(p)|$, but equality will follow.)

Sisterhood Theorem - Polygamous Case

Definition (Improvement)

A woman p is said to be *weakly better off* if:

1. $N(p)$ contains no-one who is blacklisted by p.
2. $|N(p)| \geq |O(p)|$. (*The theorem will imply equality here.)
3. For each $1 \leq i \leq O(p)$, p weakly prefers n_i^p over o_i^p.

Definition (Worsening)

A person p is said to have *gained only worse matches* if p prefers every member of $O(p)$ over every member of $N(p) \setminus O(p)$. (*Does not require $|N(p)| \leq |O(p)|$, but equality will follow.)

Theorem (Sisterhood)

If all lying women are weakly better off, then:

1. *All women are weakly better off.*
2. *All men have gained only worse matches.*

A Few Sample Corollaries

A Few Sample Corollaries

A Rural-Hospitals-type Theorem

Under the above conditions,

1. $|N(p)| = |O(p)|$ for each person $p \in W \cup M$.
2. For an innocent person p, if $|N(p)| < n_p$, then $N(p) = O(p)$.

A Few Sample Corollaries

A Rural-Hospitals-type Theorem

Under the above conditions,

1. $|N(p)| = |O(p)|$ for each person $p \in W \cup M$.
2. For an innocent person p, if $|N(p)| < n_p$, then $N(p) = O(p)$.

Corollary

If $|L| = 1$, and the lying woman is (strictly) better off, then so is some innocent woman.

A Few Sample Corollaries

A Rural-Hospitals-type Theorem

Under the above conditions,

1. $|N(p)| = |O(p)|$ for each person $p \in W \cup M$.
2. For an innocent person p, if $|N(p)| < n_p$, then $N(p) = O(p)$.

Corollary

If $|L| = 1$, and the lying woman is (strictly) better off, then so is some innocent woman.

Corollary

If all women have the same order of preference, then under the above conditions the matching must remain unchanged. Therefore, in this case there is no "significant" incentive for any subset of women to lie, even for the sake of one of them.

Questions?

Thank you!