# Stable Marriage

Suppose five men ($A$, $B$, $C$, $D$, $E$) and five women (1, 2, 3, 4, 5) enter a matchmaking service.[1]

The men and women have preferences:

| Men | Preferences |
|---|---|
| $A$ | $5 > 4 > 3 > 1 > 2$ |
| $B$ | $3 > 4 > 1 > 5 > 2$ |
| $C$ | $5 > 4 > 1 > 3 > 2$ |
| $D$ | $5 > 2 > 3 > 1 > 4$ |
| $E$ | $5 > 1 > 3 > 2 > 4$ |

| Women | Preferences |
|---|---|
| 1 | $B > D > C > A > E$ |
| 2 | $B > D > A > E > C$ |
| 3 | $B > A > D > E > C$ |
| 4 | $B > C > E > A > D$ |
| 5 | $D > B > E > A > C$ |

How should we match the men and women?

[1] Apologies for the heteronormativity.

# Rogue Couples

Suppose that we pair together $(A, 1)$ and $(B, 2)$.

- What happens if $A$ likes 2 better than his current partner, and 2 likes $A$ better than her current partner?
- There is an incentive to cheat, of course!

This is called a **rogue couple**.

A **matching** is when we pair each man with a unique woman.

The matching is called **stable** when there are no rogue couples.

Question for today: How can we find a stable matching? Do stable matchings even exist?

# Stable Matching Example

| Men | Preferences |
|---|---|
| $A$ | $5 > 4 > 3 > 1 > 2$ |
| $B$ | $3 > 4 > 1 > 5 > 2$ |
| $C$ | $5 > 4 > 1 > 3 > 2$ |
| $D$ | $5 > 2 > 3 > 1 > 4$ |
| $E$ | $5 > 1 > 3 > 2 > 4$ |

| Women | Preferences |
|---|---|
| 1 | $B > D > C > A > E$ |
| 2 | $B > D > A > E > C$ |
| 3 | $B > A > D > E > C$ |
| 4 | $B > C > E > A > D$ |
| 5 | $D > B > E > A > C$ |

Is the matching $(A, 1)$, $(B, 2)$, $(C, 3)$, $(D, 4)$, $(E, 5)$ stable?

- Do you spot any rogue couples?
- One example of a rogue couple: $(D, 5)$.

# $2 \times 2$ Stable Matchings

Men preferences match women preferences perfectly:

| Men | Preferences | Women | Preferences |
|---|---|---|---|
| $A$ | $1 > 2$ | 1 | $A > B$ |
| $B$ | $2 > 1$ | 2 | $B > A$ |

There is a unique stable matching: $(A, 1)$, $(B, 2)$.

Men have the same preferences:

| Men | Preferences | Women | Preferences |
|---|---|---|---|
| $A$ | $1 > 2$ | 1 | $A > B$ |
| $B$ | $1 > 2$ | 2 | $B > A$ |

$(A, 1)$ and $(B, 2)$ is stable, since $A$ and 1 are happy.

- Observation: If a man and woman both like each other best, they must be together in any stable matching.

# $2 \times 2$ Stable Matchings

Men and women preferences clash:

| Men | Preferences |
|---|---|
| $A$ | $1 > 2$ |
| $B$ | $2 > 1$ |

| Women | Preferences |
|---|---|
| 1 | $B > A$ |
| 2 | $A > B$ |

Here, there are two stable pairings:

- $(A, 1)$, $(B, 2)$. Stable because the men are happy.
- $(A, 2)$, $(B, 1)$. Stable because the women are happy.

So, there may be multiple stable matchings. But, so far we have always been able to find at least one stable matching.

# Stable Roommates

It is not obvious that stable matchings always exist.

Consider a variant: in the **stable roommates** problem we no longer have men/women.

| Person | Preference |
|---|---|
| $A$ | $B > C > D$ |
| $B$ | $C > A > D$ |
| $C$ | $A > B > D$ |
| $D$ | $A > B > C$ |

No matter how we assign roommates, we have a rogue pair.

- $(A, B), (C, D)$: $(B, C)$ is a rogue pair.
- $(A, C), (B, D)$: $(A, B)$ is a rogue pair.
- $(A, D), (B, C)$: $(A, C)$ is a rogue pair.

# Gale-Shapley Algorithm

Gale and Shapley gave an algorithm for finding a stable matching. Their work led to a Nobel Prize in Economics.

Since then, the algorithm has found many applications:

- Match new doctors to hospital residency programs.
- Match organ transplant patients to organs.
- And more...

Here is the algorithm (men propose version).

- On each "day", each man without a partner proposes to the women highest up in his list.
- At the end of each "day", each woman *tentatively* accepts her most preferred suitor and rejects every other man.
- Terminate when every woman has a suitor.

# Algorithm Example, Day 1

| Men | Preferences |
|---|---|
| $A$ | $5 > 4 > 3 > 1 > 2$ |
| $B$ | $3 > 4 > 1 > 5 > 2$ |
| $C$ | $5 > 4 > 1 > 3 > 2$ |
| $D$ | $5 > 2 > 3 > 1 > 4$ |
| $E$ | $5 > 1 > 3 > 2 > 4$ |

| Women | Preferences |
|---|---|
| 1 | $B > D > C > A > E$ |
| 2 | $B > D > A > E > C$ |
| 3 | $B > A > D > E > C$ |
| 4 | $B > C > E > A > D$ |
| 5 | $D > B > E > A > C$ |

Men propose:

- $B$ proposes to 3.
- $A$, $C$, $D$, and $E$ propose to 5.

Women respond:

- 3 tentatively accepts $B$.
- 5 tentatively accepts $D$; rejects $A$, $C$, $E$.

# Algorithm Example, Day 2

| Men | Preferences |
|---|---|
| $A$ | $\not{5} > 4 > 3 > 1 > 2$ |
| $B$ | $3 > 4 > 1 > 5 > 2$ |
| $C$ | $\not{5} > 4 > 1 > 3 > 2$ |
| $D$ | $5 > 2 > 3 > 1 > 4$ |
| $E$ | $\not{5} > 1 > 3 > 2 > 4$ |

| Women | Preferences |
|---|---|
| 1 | $B > D > C > A > E$ |
| 2 | $B > D > A > E > C$ |
| 3 | $\mathbf{B} > A > D > E > C$ |
| 4 | $B > C > E > A > D$ |
| 5 | $\mathbf{D} > B > E > A > C$ |

Men propose:

- $E$ proposes to 1.
- $A$ and $C$ propose to 4.

Women respond:

- 1 tentatively accepts $E$.
- 4 tentatively accepts $C$; rejects $A$.

# Algorithm Example, Day 3

| Men | Preferences |
|---|---|
| $A$ | ~~5~~ > ~~4~~ > 3 > 1 > 2 |
| $B$ | 3 > 4 > 1 > 5 > 2 |
| $C$ | ~~5~~ > 4 > 1 > 3 > 2 |
| $D$ | 5 > 2 > 3 > 1 > 4 |
| $E$ | ~~5~~ > 1 > 3 > 2 > 4 |

| Women | Preferences |
|---|---|
| 1 | $B > D > C > A >$ **E** |
| 2 | $B > D > A > E > C$ |
| 3 | **B** $> A > D > E > C$ |
| 4 | $B >$ **C** $> E > A > D$ |
| 5 | **D** $> B > E > A > C$ |

Men propose:

- $A$ proposes to 3.

Women respond:

- 3 rejects $A$ in favor of $B$.

# Algorithm Example, Day 4

| Men | Preferences |
|---|---|
| $A$ | $\not{5} > \not{4} > \not{3} > 1 > 2$ |
| $B$ | $3 > 4 > 1 > 5 > 2$ |
| $C$ | $\not{5} > 4 > 1 > 3 > 2$ |
| $D$ | $5 > 2 > 3 > 1 > 4$ |
| $E$ | $\not{5} > 1 > 3 > 2 > 4$ |

| Women | Preferences |
|---|---|
| 1 | $B > D > C > A > \mathbf{E}$ |
| 2 | $B > D > A > E > C$ |
| 3 | $\mathbf{B} > A > D > E > C$ |
| 4 | $B > \mathbf{C} > E > A > D$ |
| 5 | $\mathbf{D} > B > E > A > C$ |

Men propose:

- *A* proposes to 1.

Women respond:

- 1 rejects *E* in favor of *A*.

# Algorithm Example, Day 5

| Men | Preferences |
|---|---|
| *A* | ~~5~~ > ~~4~~ > ~~3~~ > 1 > 2 |
| *B* | 3 > 4 > 1 > 5 > 2 |
| *C* | ~~5~~ > 4 > 1 > 3 > 2 |
| *D* | 5 > 2 > 3 > 1 > 4 |
| *E* | ~~5~~ > ~~1~~ > 3 > 2 > 4 |

| Women | Preferences |
|---|---|
| 1 | *B* > *D* > *C* > **A** > *E* |
| 2 | *B* > *D* > *A* > *E* > *C* |
| 3 | **B** > *A* > *D* > *E* > *C* |
| 4 | *B* > **C** > *E* > *A* > *D* |
| 5 | **D** > *B* > *E* > *A* > *C* |

Men propose:

- *E* proposes to 3.

Women respond:

- 3 rejects *E* in favor of *B*.

# Algorithm Example, Day 6

| Men | Preferences |
|---|---|
| *A* | ~~5~~ > ~~4~~ > ~~3~~ > 1 > 2 |
| *B* | 3 > 4 > 1 > 5 > 2 |
| *C* | ~~5~~ > 4 > 1 > 3 > 2 |
| *D* | 5 > 2 > 3 > 1 > 4 |
| *E* | ~~5~~ > ~~1~~ > ~~3~~ > 2 > 4 |

| Women | Preferences |
|---|---|
| 1 | *B* > *D* > *C* > **A** > *E* |
| 2 | *B* > *D* > *A* > *E* > *C* |
| 3 | **B** > *A* > *D* > *E* > *C* |
| 4 | *B* > **C** > *E* > *A* > *D* |
| 5 | **D** > *B* > *E* > *A* > *C* |

Men propose:

- *E* proposes to 2.

Women respond:

- 2 tentatively accepts *E*.

Since all women now have a suitor, the algorithm terminates with the pairing $(A, 1)$, $(B, 3)$, $(C, 4)$, $(D, 5)$, $(E, 2)$.

# Algorithm Termination

**Theorem**: The Gale-Shapley algorithm terminates in finite time for any stable marriage instance.

*Proof*.

- Each day of the algorithm, at least one man is proposing.
- If no man gets rejected, the algorithm immediately terminates.
- So, each day that the algorithm runs, at least one man gets rejected.
- So, each day that the algorithm runs, at least one man crosses off one woman from his preference list.
- If there are $n$ men, the men's preference lists have a total of $n^2$ entries, so the algorithm terminates in $\leq n^2$ days. □

# Algorithm Termination

If every man has a different first choice, how many days does the algorithm require? Just one!

On the other hand, there *are* instances which require $\Omega(n^2)$ days to complete, where $n$ is the number of men/women.

There can be *exponentially many* stable marriages.

# Algorithm Correctness

**Theorem**: The Gale-Shapley algorithm output is stable.

*Proof.*

- Suppose the result of the algorithm is not stable, i.e., there is a rogue couple $(M, W)$.
- Say that the algorithm pairs $(M, W')$ and $(W, M')$.
- So, $M$ prefers $W$ over $W'$, and $W$ prefers $M$ over $M'$.
- So, $M$ proposes to $W$ before he proposes to $W'$.
- But, since $M$ proposes to $W'$ later, that means he must have gotten rejected by $W$.
- That means $W$ must have found a better guy than $M$. But $W$'s suitors can only get better each day!
- This means $W$ must end up with a guy she likes better than $M$. Contradiction. □

# Which Stable Matching?

We saw that there can be multiple stable matchings. Which one does the Gale-Shapley algorithm output?

| Men | Preferences |
|---|---|
| $A$ | $1 > 2$ |
| $B$ | $2 > 1$ |

| Women | Preferences |
|---|---|
| 1 | $B > A$ |
| 2 | $A > B$ |

Recall: The stable matching $(A, 1)$, $(B, 2)$ favors the guys, and $(A, 2)$, $(B, 1)$ favors the gals.

Run Gale-Shapley. The algorithm ends in one day, and the pairing is $(A, 1)$, $(B, 2)$.

Does the Gale-Shapley algorithm favor the men?

# Optimality

We say that a stable matching is **optimal** for a man if his partner in this matching is the best possible partner he can have, *out of all possible stable matchings*.

| Men | Preferences |
|---|---|
| $A$ | $1 > 2$ |
| $B$ | $2 > 1$ |

| Women | Preferences |
|---|---|
| 1 | $B > A$ |
| 2 | $A > B$ |

Which stable matchings are optimal for $A$?

- All possible stable matchings: $(A,1),(B,2)$ and $(A,2),(B,1)$.
- What are all the possible women that $A$ can end up with, *out of the possible stable matchings*? $\{1,2\}$
- Out of these women, which does $A$ like best? 1
- Therefore, any stable matching in which $A$ ends up with 1 is optimal for $A$.

# Optimality vs. Most Preferred

Is a man's optimal partner the same as the first person in his preference list?

| Men | Preferences |
|---|---|
| $A$ | $1 > 2$ |
| $B$ | $1 > 2$ |

| Women | Preferences |
|---|---|
| 1 | $B > A$ |
| 2 | $B > A$ |

What are the stable matchings? Only $(A, 2), (B, 1)$.

In *any* stable matching, $A$ must be partnered with 2. So the best he can do is 2 (his optimal partner).

So, $A$'s optimal partner is *not* the same as his most preferred partner. 1 is unattainable for $A$.

# Male Optimality

We say that a stable matching is **male optimal** if *every* man is paired with his optimal partner.

- This is a pretty strong condition— all men are happy simultaneously!

| Men | Preferences |
|---|---|
| $A$ | $1 > 2$ |
| $B$ | $2 > 1$ |

| Women | Preferences |
|---|---|
| 1 | $B > A$ |
| 2 | $A > B$ |

Here, $(A, 1), (B, 2)$ is male optimal because it is stable and every man gets his first choice.

# Male Optimality for Gale-Shapley

**Theorem**: The Gale-Shapley algorithm outputs a male optimal stable matching.

*Proof*.

- Suppose that the output is not male optimal.
- Consider the first day in which some man $M$ is rejected by his optimal partner $W$ (Well Ordering Principle).
- Since $W$ rejects $M$, that means she had a man she likes better: call him $M'$.
- $M'$ has not yet been rejected by his optimal woman. So, $W$ is at least as good as the optimal woman for $M'$.
- Since $W$ is optimal for $M$, there exists a stable matching in which $M$ is paired with $W$.
- But $W$ likes $M'$ more than $M$, and $M'$ likes $W$ at least as much as his partner. This is a rogue couple. □

# Notes on the Optimality Result

How do we modify the Gale-Shapley algorithm to favor the women?

- Switch the roles of men and women.
- In other words, the women propose.

The Gale-Shapley algorithm is good for men. Is it necessarily bad for women?

- A man is **pessimal** for a woman if, out of all stable matchings, this man is her least preferred partner.
- A stable matching is **female pessimal** if every woman is with her pessimal partner.

# Male Optimal Is Female Pessimal

**Theorem**: Consider two different stable matchings, $\mu_1$ and $\mu_2$. If every man likes $\mu_1$ as much as $\mu_2$, then every woman likes $\mu_2$ as much as $\mu_1$.

*Proof.*

- Take a pair $(M, W)$ in $\mu_1$ which is not matched in $\mu_2$.
- If both $M$ and $W$ prefer $\mu_1$, that means $M$ and $W$ like each other more than their partners in $\mu_2$.
- This is a rogue couple!
- So, if $M$ prefers $\mu_1$, then $W$ prefers $\mu_2$. □

Every man prefers the male optimal matching over any other stable matching. So, every woman prefers any other stable matching over the male optimal matching.

- The male optimal matching is female pessimal.

# Male Optimal Is Female Pessimal

More details:

- Every woman prefers any other stable matching over the male optimal matching.
- For a particular woman $W$, let her pessimal partner be $M$.
- There is some stable matching in which $M$ and $W$ are paired together.
- Since $W$ prefers this matching over the male optimal matching, that means $W$ must be matched with $M$ in the male optimal matching.
- So, every woman is matched with her pessimal partner in the male optimal matching.

# Summary

- We are given $n$ men and $n$ women with preference lists. We want a matching: an assignment of men to women.
- A rogue couple is a pair who prefer each other to their partners. A matching without a rogue couple is stable.
- The Gale-Shapley algorithm outputs a stable matching.
- A stable matching is male optimal if every man prefers this matching over any other stable matching.
- A stable matching is female pessimal if every woman prefers any other stable matching over this matching.
- The Gale-Shapley algorithm (with men proposing) is male optimal and female pessimal.