Networks II: Market Design—Lecture 16 Matching Markets with Non-transferable Utilities: Two-sided preferences

ARPITA GHOSH

Dept. Of Information Science,
Cornell University

What we've seen so far

- Stable matchings in the marriage model: Constructive proof of existence
  - Deferred acceptance (DA) algorithm returns stable matching
  - Argument assuming (strict) complete preferences, $|M| = |W|$
  - Extend model, algorithm: $|M| \neq |W|$; (strict) incomplete preferences
  - Incomplete preferences: Stability means no blocking pairs **and** individual rationality
- Deferred acceptance produces **a** stable matching: Other stable matchings may exist!
- Next: Comparing matchings

Comparing matchings

For any two matchings μ and μ', write

- $\mu \succeq_m \mu'$ if m (weakly) prefers his partner in μ to μ'
  (Either $\mu(m) \succ_m \mu'(m)$ or $\mu(m) = \mu'(m)$)
- If $\mu \succeq_m \mu'$ for **all** $m \in M$: Say $\mu \succeq_M \mu'$
  - **All** men prefer matching μ to matching μ'
- Similar definitions for each woman w, and set of women W:
  - $\mu \succeq_w \mu'$ if if w (weakly) prefers her partner in μ to μ'
    (Either $\mu(w) \succ_w \mu'(w)$ or $\mu(w) = \mu'(w)$)
  - If $\mu \succeq_w \mu'$ for **all** $w \in W$: $\mu \succeq_W \mu'$
    - **All** women prefer matching μ to matching μ'

An example

- Consider a market with preference profiles:

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

- Consider two matchings:
  - $\mu_1 = (m_1 - w_1, m_2 - w_2, m_3 - w_3)$
  - $\mu_2 = (m_1 - w_2, m_2 - w_3, m_3 - w_1)$
- Which is true? [A] $\mu_1 \succeq_M \mu_2$ [B] $\mu_1 \succeq_W \mu_2$ [C] Both [D] Neither
- Which agents prefer matching μ_1 to matching μ_2?
  - m_1, m_3, w_3 prefer μ_1
  - Remaining prefer μ_2

An example

- A different set of preference profiles:

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

- Again, consider the two matchings:
  - $\mu_1 = (m_1 - w_1, m_2 - w_2, m_3 - w_3)$
  - $\mu_2 = (m_1 - w_2, m_2 - w_3, m_3 - w_1)$
- Which is true? [A] $\mu_2 \succeq_M \mu_1$ [B] $\mu_2 \succeq_W \mu_1$ [C] Both [D] Neither
  - Note that μ_2 gives both men and women their top choices
  - (μ_2 is outcome of both man and woman proposing DA)

Comparing matchings

- Given any two matchings μ_1 and μ_2 (and strict preferences):
  - Any *single* agent $a \in M \cup W$ can compare μ_1 to μ_2: Either $\mu_1 \succeq_a \mu_2$ or vice versa
  - But not all agents on one side need agree about comparison
- Suppose $\mu_1 \succeq_M \mu_2$: Which of the following must hold?
  - $\mu_1 \succeq_W \mu_2$
  - $\mu_2 \succeq_W \mu_1$
  - Either $\mu_1 \succeq_W \mu_2$ or $\mu_2 \succeq_W \mu_1$
  - None of the above
  - All men agreeing does not imply all women will also agree about which matching is better (E.g.: m_i's top choice is w_i, all women prefer m_1)
- In general, arbitrary pair of matchings need not be *comparable*: **Neither** $\mu_1 \succeq_M \mu_2$ nor $\mu_2 \succeq_M \mu_1$ may hold

Back to our example

- Consider market with preferences:

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

- What is outcome of DA with men proposing?
  - $\mu_M = (m_1 - w_2, m_2 - w_1, m_3 - w_3)$
- What is outcome of DA with women proposing?
  - $\mu_W = (m_1 - w_3, m_2 - w_1, m_3 - w_2)$
- Are matchings μ_M and μ_W comparable?
  - $\mu_M \succeq_M \mu_W$ and $\mu_M \succeq_W \mu_W$
  - $\mu_M \preceq_M \mu_W$ and $\mu_M \preceq_W \mu_W$
  - $\mu_M \succeq_M \mu_W$ and $\mu_M \preceq_W \mu_W$
  - None of the above

Another example

- Recall example with three stable matchings:

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

- $\mu_M = (m_1 - w_1, m_2 - w_2, m_3 - w_3), \mu_W = (m_1 - w_2, m_2 - w_3, m_3 - w_1)$, $\mu = (m_1 - w_3, m_2 - w_1, m_3 - w_2)$
- Are the matchings μ_M, μ and μ_W comparable?
  - Only μ_M and μ_W are comparable
  - Only μ comparable with both μ_W and μ_M
  - No pair of matchings is comparable
  - All three matchings are comparable
    - Yes: In fact, $\mu_M \succeq_M \mu \succeq_M \mu_W$ and $\mu_M \preceq_W \mu \preceq_W \mu_M$!

Comparing stable matchings: Some questions

- Compare outcomes $\mu_M, \ \mu_W$ of man-proposing and woman-proposing DA algorithms
- In our examples:
  (i) Matchings were comparable (!)
  (ii) Men preferred μ_M
  (iii) Women preferred μ_W
- Is this true in general?
  - Do men prefer outcome from men-proposing DA to outcome from women-proposing DA, and vice-versa?

- Second example: $\mu_M \succeq_M \mu$ for other (stable) matching μ as well
- Question: Consider set of all stable matchings
- Is it possible that there is a stable matching μ^* such that $\mu^* \succeq_M \mu$ for *all* stable matchings μ?
  (And similarly for women?)

Optimal stable matchings

Theorem (Gale and Shapley 1962)

- *There exists a man-optimal stable matching, that is, a stable matching that every man weakly prefers to every other stable matching (and similarly for women).*
- *The man-proposing DA algorithm returns this man-optimal matching (and the woman-proposing DA returns the woman-optimal matching).*

- Men agree on which is the best *stable* matching, *even though* competing amongst each other!
- Man-proposing DA returns this matching, even though men's partners get progressively worse through the course of the algorithm!

Proving the result

Theorem: Men-proposing DA returns man-optimal stable matching

Structure of the argument:

- Define notion of achievability: A partner you can feasibly 'hope to get'
- Lemma: No man is rejected by his most-preferred achievable woman in man-proposing DA algorithm
- Finishing the proof:
  - No man proposes to a woman ranked below his most-preferred achievable women in DA (from Lemma; definition of algorithm)
  - No man is matched by DA to a higher-ranked woman (Why?)
  - Every man is matched to top achievable woman: DA returns man-optimal stable matching

Achievability

- Achievable partner: w is **achievable** for m if there is some **stable** matching μ where $\mu(m) = w$ (and similarly for women)

- 3 men and 3 women with preferences:

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

- Is w_2 achievable for m_1? ([A] Yes [B] No)
  - No: m_2 and w_2 are each other's top choices

- Is m_1 achievable for w_3? ([A] Yes [B] No)
  - Yes: Woman-proposing DA

Men-proposing DA returns man-optimal stable matching

- Lemma: No man is rejected by his top-ranked achievable woman in man-proposing DA algorithm
  - Achievable women for m: Set of m's partners across all stable matchings
- Suppose not: Consider *first* round in which a man, say m, is rejected by an *achievable* woman, say w
  - This means that some other man m' proposed to w and was tentatively accepted by w in this step
  - w is achievable for m: Let $\hat{\mu}$ be a stable matching where $\hat{\mu}(m) = w$
  - Does m' like w or $w' = \hat{\mu}(m')$ better?
    - m' proposes to w in DA: Before or after proposing to $\hat{\mu}(m')$?
    - If after: m' must have rejected by $\hat{\mu}(m')$
    - But m is *first* to be rejected by an achievable woman
    - So m' must have proposed to w before $\hat{\mu}(m')$ in DA: m' must prefer w to $w' = \hat{\mu}(m')$
  - Also w prefers m' to m (since she rejects m for m' in DA)
- So pair (m', w) blocks $\hat{\mu}$: Contradiction to stability of $\hat{\mu}$!

Appreciating the result

- Consider set S of all stable matchings μ: w is achievable for m if $w = \mu(m)$ for some $\mu \in S$
- Every man m points to favorite partner $w^*(m) = \mu(m)$ over all matchings μ in S
- Consider pairs $(m, w^*(m))$ for all $m \in M$: (Man, best achievable woman) pairs
- Result says:
  - This set of pairs forms a matching (!!)
  - This matching is in fact stable
  - It is exactly the matching returned by the man-proposing DA algorithm!
- (Similarly for women)