Networks II: Market Design—Lecture 6 Matching Markets with Non-transferable Utilities: One-sided rank-order preferences

ARPITA GHOSH

Dept. Of Information Science,
Cornell University

Recap: Last time

- Mechanisms
- Good mechanisms: Pareto-efficiency
  - Mechanism φ is **Pareto-efficient** if it returns a Pareto-efficient matching $M = \varphi(A, H, \succ)$ for every input market $(A, H, \succ)$
- The serial dictatorship mechanism

Outline: This class and next

- Proof of Pareto-efficiency of serial dictatorship
- Incentives and preference reporting: Strategy-proofness
- Strategyproofness of serial dictatorship
- How general are serial dictatorships?

Recall: The serial dictatorship mechanism

- Serial dictatorship, or *priority* mechanisms:
  - Serial dictatorship mechanism specifies priority order f over A: $f(i)$ is agent with i^{th} priority
  - Given market $(A, H, \succ)$: Agent $f(i)$ receives item ranked highest in her preference list amongst all items except those already taken by agents $f(1), \ldots, f(i-1)$

Theorem

The serial dictatorship mechanism is Pareto-efficient when agents have strict rank-order preferences over items.

Pareto-efficiency of serial dictatorship

Why?

- Assume for now that all houses are *acceptable* to all agents
- Suppose serial dictatorship mechanism returns a matching $M = \varphi(\succ)$ in some market $(A, H, \succ)$ that is Pareto-dominated by matching $\hat{M}$
- Consider *highest-priority* agent who obtains different houses in M and $\hat{M}$, say agent a: $M(a) \neq \hat{M}(a)$
- Two facts:
  - a cannot prefer $M(a)$ to $\hat{M}(a)$: So $\hat{M}(a) \succ_a M(a)$ (Why?) (Note: We are using strict preferences assumption here!)
  - a cannot prefer $\hat{M}(a)$ to $M(a)$: So $M(a) \succ_a \hat{M}(a)$ (Why?)
- Contradiction: Matching M must have been Pareto-efficient!

Pareto-efficiency and input preferences

- Mechanism φ is **Pareto-efficient** if it returns a Pareto-efficient matching $M = \varphi(A, H, \succ)$ for every **input** market $(A, H, \succ)$
- Suppose a Pareto-efficient mechanism uses input $\succ'$ instead of *true* preferences $\succ$
- Output matching $M = \varphi(\succ')$ is Pareto-efficient with respect to $\succ'$, but not necessarily with respect to $\succ$!
- Example:
  - Matching $M =$ (a-x, b-z, c-y) wrt to preferences $\succ$:
    a: x $\succ$ z $\succ$ y; b: x $\succ$ z $\succ$ y; c: x $\succ$ y $\succ$ z
    is Pareto-efficient, but *not* wrt preferences
    $\succ'$ a: z $\succ$ x $\succ$ y; b: x $\succ$ z $\succ$ y; c: x $\succ$ y $\succ$ z

Pareto-efficiency and input preferences

- Where do the input preferences $\succ$ come from?
  - Mechanism: 'Algorithm with input from agents'
  - Input preferences $\succ$: Reported by agents
  - Agents may not always want to report true preferences!
- Pareto-efficiency of output matching w.r.t **true** preferences depends on *what agents report to mechanism*

Why care?

- Consider two mechanisms, and market with 2 agents a,b; 2 houses x,y
  - Preference profile $\succ^1$: $\succ^1_a, \succ^1_b$: $x \succ y$
  - Preference profile $\succ^2$: $\succ^2_a = x \succ y$; $\quad \succ^2_b = x$
- Mechanism φ_1 returns the following matchings:
  - $\varphi_1(\succ^1) =$ (a-x,b-y); $\varphi_1(\succ^2) =$ (a-x)
- Mechanism φ_2 returns the following matchings:
  - $\varphi_2(\succ^1) =$ (a-x,b-y); $\varphi_2(\succ^2) =$ (a-y, b-x)
- 'How bad' is it to use the 'wrong' input?
  - φ_1: Using $\succ^2$ instead of $\succ^1$ harms only *b*
  - φ_2: Using $\succ^2$ instead of $\succ^1$ harms *a* and helps *b*!

Strategyproofness

- Recall: (Deterministic) mechanism assigns a matching for each preference profile
  - $\varphi(\succ)$: Matching when agents report $\succ$ under mechanism φ
- Agents' *strategies*: What preference profile $\succ_i$ to report
- **Strategyproofness**: Mechanism φ is **strategyproof** if agents cannot benefit by lying about their preferences, *no matter what other agents report*

Strategyproofness, formally

- Mechanism φ returns matching $M = \varphi(\succ)$; all agents **know** mechanism
  - Consider agent a with true preferences $\succ_a$
  - Other agents report preferences $\succ_{-a}$
  - $M = \varphi(\succ_a, \succ_{-a})$; $M' = \varphi(\succ'_a, \succ_{-a})$
- Mechanism φ is strategyproof if $M(a) \succeq_a M'(a)$ for all $\succ_{-a}$, for all agents a (and corresponding preferences $\succ_a$)
  - Does it matter whether remaining agents report true preferences? ([A] Yes [B] No)
  - Does it matter whether an agent can see other agents' reports? ([A] Yes [B] No)

Strategyproofness, formally

- Mechanism φ returns matching $M = \varphi(\succ)$; all agents **know** mechanism
  - Consider agent a with true preferences $\succ_a$
  - Other agents report preferences $\succ_{-a}$
  - $M = \varphi(\succ_a, \succ_{-a})$; $M' = \varphi(\succ'_a, \succ_{-a})$
- Mechanism φ is strategyproof if $M(a) \succeq_a M'(a)$ **for all** $\succ_{-a}$, for all agents a (and corresponding preferences) $\succ_a$
  - Does it matter whether remaining agents report true preferences?
  - Does it matter whether an agent can see other agents' reports?
  - No, and no: Notice "**for all** $\succ_{-a}$" in definition!
- Strategyproofness: Strong solution concept (cf. Nash equilibrium)

Strategyproofness: Revisiting an example

- Example: Agents a,b,c; houses x,y,z
- Preferences $\succ$ are a: x $\succ$ z $\succ$ y; b: x $\succ$ z $\succ$ y; c: x $\succ$y $\succ$ z
- $\varphi(\succ)$ returns matching a-y, b-x, c-z: Is φ strategyproof?
  - A Yes
  - B No
  - C How would I know?
- While we're at it: Is φ Pareto-efficient?
  - A Yes
  - B No
  - C How would I know?
- Note that strategyproofness is:
  - A property of a *mechanism*, not of a matching
  - *Not the same* as Pareto-efficiency!

Strategyproofness: Another example

- Consider market with agents a,b,c; houses x,y,z
  - Preferences $\succ_1$ are a: $x \succ y \succ z$; b: $x \succ y \succ z$; c: $x \succ y \succ z$
  - Preferences $\succ_2$ are a: $x \succ y \succ z$; b: $x \succ y \succ z$; c: $y \succ x \succ z$
  - $\varphi(\succ_1)$: Matching a-x, b-y, c-z; $\varphi(\succ_2)$: Matching a-x, b-z, c-y
- Is φ strategyproof?
  - A Yes
  - B No
  - C Not again—you know I can't know!
- Suppose actual preferences are $\succ_1$: Can any agent benefit by lying about her preferences?

Strategyproofness: Continuing with this example

- Let's try generalizing what happened in previous example:
  - Consider two input preference profiles $\succ$ and $\succ'$ that differ only in agent i's preference: $\succ_i \neq \succ'_i$
  - Suppose agent i's allocations with two inputs are *different*:
    - Allocation h in $\varphi(\succ)$, h' in $\varphi(\succ')$
    - $h \neq h'$
  - Preferences are strict: So either $h \succ h'$ or $h' \succ h$

- So can we conclude: Agent i can benefit from lying in one of $\succ$ or $\succ'$— *i.e.*, φ is not strategyproof?
  [A] Yes [B] No
  - No—we can't!
  - i needs to prefer h' to h in ranking $\succ$, **or** h to h' in $\succ'$, to demonstrate strategyproofness, *i.e.*, benefit from lying
    - If $h \succ_i h'$ and $h' \succ'_i h$, lying does not benefit i for true preference being either $\succ_i$ or $\succ'_i$

Yet another example

- Recall assumption of *complete* preferences (*i.e.*, all agents find all houses acceptable)
- *'Arbitrary assignment'* mechanism:
  - $|A| = |H| = n$: $n!$ possible perfect matchings; all are feasible
  - Mechanism always returns a certain perfect matching M
- Is this mechanism (i) strategyproof? (ii) Pareto-efficient?
  - A Only strategyproof
  - B Only Pareto-efficient
  - C Both strategyproof and Pareto-efficient
  - D Neither strategyproof nor Pareto-efficient

- Pareto-efficiency (PE): Mechanism $\varphi(\succ)$ is PE with respect to *reported* preferences
- Strategyproofness (SP): Mechanism $\varphi(\succ)$ elicits *true* preferences
- Mechanism that is both SP and PE: Pareto-efficient with respect to *true* preferences