Top Trading Cycles Mechanism

Qianfeng Tang (SUFE)

Summer School of Modern Economics
Nanchang University, July 2015

House allocation

Objects (houses) $H = \{h_1, h_2, \ldots, h_n\}$ are to be allocated to agents $N = \{1, \ldots, n\}$:

- Each object has exactly one copy
- Each agent i has a strict preference P_i over H; $P \equiv (P_i)_{i \in N}$
- A matching (allocation) is a bijection $\mu : N \to H$
- No monetary transfer allowed

Applications include kidney exchange, allocation of housing/offices, etc.

Matching criteria

A mechanism φ selects a matching $\varphi(P)$ for each preference profile P; use $\varphi_i(P)$ to denote i's assignment at P

When allocating objects to agents, certain properties are desirable:

- A matching μ is **Pareto efficient** (for students) if there is no matching $\nu \neq \mu$ such that $\nu(i) R_i \mu(i), \forall i$ and $\nu(j) P_j \mu(j)$ for some agent j
- φ is **Pareto efficient** if it selects a Pareto efficient matching for all P
- φ is **strategy-proof** if no student can benefit from misreporting. Formally, if $\varphi_i(P_i, P_{-i}) R_i \varphi_i(P'_i, P_{-i})$, for all i, P'_i, P_{-i}.
- φ is **group strategy-proof** if there do not exist $C \subset I$ and P'_C such that for all $i \in C, \varphi_i(P_C, P_{-C}) R_i \varphi_i(P'_C, P_{-C})$, and for some $j \in C, \varphi_j(P_C, P_{-C}) P_j \varphi_j(P'_C, P_{-C})$.

Serial dictatorship

Let's begin with the simplest "good" mechanism

Serial dictatorship: An order, which is a bijection $f : N \to N$, is given *a priori*. The first agent (in the order) chooses her favorite object first, then the second agent chooses her favorite from what remain, and so on.

Example

Suppose $H = \{a, b, c\}, N = \{1, 2, 3\}$, and agents have the following preferences:

1	2	3
a	a	b
b	c	c
c	b	a

Serial dictatorship

Let r be a relabeling (permutation) of houses and P^r be the new preferences. A mechanism φ is **neutral** (to the names of objects) if $\varphi(P^r) = r(\varphi(P))$.

DA with non-uniform priorities is not neutral, and DA with uniform priority is exactly SD with this priority as serial order.

SD is Pareto efficient, group strategy-proof, and neutral. Furthermore,

Theorem (Svensson, 1999)

A mechanism is group strategy-proof and neutral if and only if it is a serial dictatorship.

The necessity part is not difficult. See Svensson's paper for proof of the sufficiency part

Shapley-Scarf housing market

A **housing market** is a profile $\{(i, h_i)_{i \in N}, P\}$ such that:

- agent i owns house $h_i, \forall i$
- agent i has strict preference P_i over H

Gale's **top trading cycles** (TTC) mechanism:

Step 1 Each agent points to the owner of her most favorite house. Due to finiteness, there exists at least one cycle (including self-cycles) and cycles don't intersect. Let agents in cycles trade and remove them.

Step $k, k \geq 2$ Repeat Step 1 with the remaining agents until all are removed.

Example

Suppose $N = \{1, 2, 3, 4, 5\}$ and each agent i is endowed with house h_i. Consider preference profile

1	2	3	4	5
h_2	$\boxed{h_3}$	$\boxed{h_2}$	h_3	$\boxed{h_1}$
h_4	$\vdots$	$\vdots$	$\boxed{h_4}$	$\vdots$
$\boxed{h_5}$			$\vdots$	
$\vdots$				

The core

A matching μ is in the **core** of the market if there exists no coalition $B \subset N$ and matching ν such that

- For any $i \in B, \nu(i)$ is owned by some agent in B
- For agents in B, matching ν Pareto dominates μ

A core allocation is both individually rational (consider B as singleton set) and Pareto efficient (consider B as the whole set N)

Theorem (Roth and Postlewaite, 1977)

The matching produced by Gale"s TTC is the unique core matching of housing market.

Proof: Core

Let μ be the TTC assignment. Suppose instead μ is not in the core

- By definition of core, μ is blocked by some coalition B via matching ν
- Let i be the earliest matched agent in TTC among agents in B who are strictly better off under ν. Then $\nu(i)$ must have been removed earlier than i in TTC
- Suppose $\nu(i) = h_j$, the house of agent j. Then $j \in B$ and is matched earlier than i in TTC. By assumption on i, j is not strictly better off under ν, i.e., $\mu(j) = \nu(j)$
- Likewise, $j \in B$ implies that the owner of $\nu(j)$ is in B, is removed together with j in TTC, and is matched the same under μ and ν
- By induction, the agent who obtains h_j in TTC is in B, and since h_j is assigned to i at ν, this agent must be strictly improved at ν. Since she is removed earlier than i, we have a contradiction

Proof: uniqueness

Suppose other than the TTC matching μ, ν is another matching in the core

- Then for some $i, \mu(i) \neq \nu(i)$
- Note that first, i cannot be any agent who is matched in the first round of TTC, because such agents obtain their most favorite at μ and $\nu(i) \neq \mu(i)$ implies that $\nu(i)$ is worse
- Therefore, i can always block ν by forming a coaltion with members in the same cycle in TTC
- By the same argument i cannot be any agent in the second round of TTC, and so on

Ma's characterization

A matching μ is **individually rational (IR)** if $\mu(i) R_i h_i$ for all i

Theorem (Ma, 1994)

A mechanism is strategy-proof, Pareto efficient and individually rational if and only if it is TTC.

That is, for housing markets, if a mechanism always selects the IR and Pareto efficient outcome, and is strategy-proof, then it must always select the core outcome

Proof

Suppose a mechanism φ satisfies all three axioms above. Let P be any preference profile

- First, consider any agent i who trades in the *first* step of TTC. Suppose instead $\varphi_i(P) \neq TTC_i(P)$
- If i trades with herself in TTC, then since φ satisfies IR, $\varphi_i(P) = TTC_i(P) = h_i$
- Otherwise, i trades with others in TTC. For simplicity, suppose it is a two-way cycle $i \leftrightarrow j$
- Then i top ranks h_j and $TTC_i(P) = h_j$ is better than $\varphi_i(P)$. Consider an alternative preference $P_i' : h_j h_i \emptyset$ of i
- Since φ is strategy-proof, $\varphi_i(P_i', P_{-i}) \neq h_j$, because otherwise i has incentive to misreport P_i' when her true preference is P_i ; hence due to IR, $\varphi_i(P_i', P_{-i}) = h_i$
- Consequently, $\varphi_j(P_i', P_{-i}) \neq h_i$ and j is worse off than in TTC. Similarly, consider $P_j' : h_i h_j \emptyset$
- Since φ is strategy-proof, $\varphi_j(P_i', P_j', P_{-ij}) \neq h_i$ and due to IR, $\varphi_j(P_i', P_j', P_{-ij}) = h_j$
- At the preference profile $(P_i', P_j', P_{-ij}), \varphi$ is not Pareto efficient, a contradiction. The rest of the proof follows from induction.

Application: Kidney exchange

Some background:

- National Organ Transplant Act, 1984, makes it illegal to buy or sell kidney
- over **90,000** patients are on waitlists for kidneys in the U.S. (with only around **17,000** donors in total)
- In 2006, there are 10659 transplants from disceased donors; 6428 transplants from live donors, and
- In a recent year, **4500** patients died while on the waiting list

- transplant from live donor survives much longer than from deceased donor
- issue of relative or friend donation: blood-type incompatibility or antibodies
- in 1986, the idea of **live-donor paired kidney exchanges** is proposed
- by 2005, only 10% of of live donor transplants are from live-donor exchanges

House allocation with existing tenants

Houses in H_O are occupied by existing tenants and houses in H_V are vacant. Application: Dormitory allocation and kidney exchange.

The **YRMH-IGYT** algorithm (Abdulkadiroglu and Sonmez, 1999) combines SD and TTC. Before it starts, a priority order f of agents is exogenously given or randomly drawn.

Equivalently, it operates as follows:

Step 1 All vacant houses point to the highest priority agent; every occupied house points to its owner. Trade and remove cycles.

Step $k, k \geq 1$ Repeat Step 1 with the remaining houses and agents

General allocation rules

By now:

- For markets with ownership (e.g., housing market): Trading mechanisms
- For markets without ownership (e.g., house allocation): Serial dictatorship
- Mixed ownership (e.g., allocation with existing tenants): Mixed mechanism

For general allocation problems

- Endow: create a housing market; and then apply
- Trading mechanism

Hierarchical exchange rules

An **inheritance tree** of an object specifies

- to whom this object is initially endowed
- depend on the assignment history of previous owners, who inherits it

A **hierarchical exchange rule** (Papai, 2000) operates as follows:

Step 1 Objects are initially endowed to agents according to its inheritance tree. Agents trade according to TTC.

Step $k, k \geq 2$ After each previous trading history, after agents are removed, remaining objects are inherited by remaining agents according to the inheritance trees. Then trade and remove as usual.

Examples

Sequential Dictatorship. Let $N = \{1, 2, 3\}$ and $H = \{a, b, c\}$. Agent 1 chooses first. And if 1 chooses a or b, then 2 chooses next; if 1 chooses c, then 3 chooses next.

TTC in School Choice. This mechanism is introduced by Abdulkadiroglu and Sonmez (2003); it extends Gale's TTC from housing markets to priority-based allocation problems. In the algorithm, the priority list of a school is interpreted as the order that the students inherit seats from the school

For each school choice problem, TTC operates as follows:

- Step 1 Each student points to her most favorite school; each school points to its highest priority student. There will be cycle(s). Match students in cycles with their favorite seats and remove them and those seats.
- Step $k, k \geq 2$ Repeat Step 1 with the remaining seats and students.

Unification of existing mechanisms

- **Serial dictatorship**: All objects are endowed to the first agent in the order and all left objects are always inherited by the next agent in the order. In trading, every agent points to herself
- **TTC**: Each object is endowed to a different agent. No inheritance happen
- **YRMH-IGYT**: All vacant houses are endowed and inherited according to the priority; each occupied house is owned by the respective existing tenant

A mechanism is **reallocation-proof** if no pair of agents can gain from misreporting and swapping ex post in a self-enforcing way.

Theorem (Papai, 2000)

An allocation mechanism is group strategy-proof, Pareto efficient, and reallocation-proof if and only if it is a hierarchical exchange rule.