CS 6002: Selected Areas of Mechanism Design **Jan-Apr 2025**

Lecture 2: Top Trading Cycle Mechanism

Lecturer: Swaprava Nath *Scribe(s): Harsh Suthar*

Disclaimer: *These notes aggregate content from several texts and have not been subjected to the usual scrutiny deserved by formal publications. If you find errors, please bring to the notice of the Instructor.*

2.1 Top Trading Cycle Mechanism

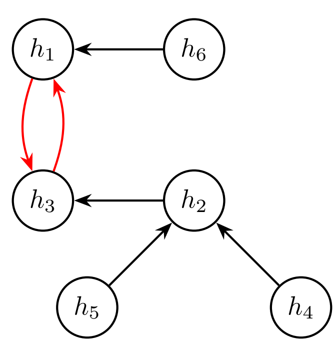
2.1.1 Initial Setup

- Suppose we have 6 people and 6 houses ($|N| = |M| = 6$).
- Each person is initially allocated a house by a third party arbitrarily, say $P_i \longrightarrow h_i \quad \forall i \in M$.
- Let the preference order of each of the 6 agent be the following:

P_1	P_2	P_3	P_4	P_5	P_6
h_3	h_3	h_1	h_2	h_2	h_1
h_1	h_2	h_4	h_1	h_1	h_3
h_2	h_1	h_3	h_5	h_6	h_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	h_4
					$\vdots$

2.1.2 Constructing the Graph

Each person P_i initially owns a house h_i. Construct a directed graph by adding a directed edge from the house own by an agent to its most preferred house in the first round.

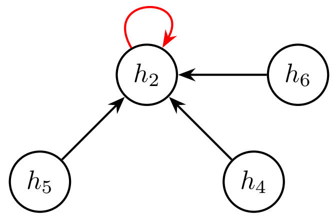


Round 1 of TTC

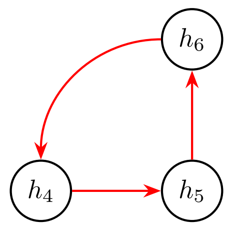
Since each node has an outgoing edge, we are guaranteed to have a cycle in the graph. In the graph it can be observed that h_1 and h_3 form a cycle. We can allocate h_3 to P_1 and h_1 to P_3, ensuring that both of them get their top priorities.

Now we delete both h_1 and h_3 from the graph and update the preferences of all the nodes that pointed to h_1 and h_3 to their next highest preference, since these two houses have already been allocated.

Round 2 of TTC



P_2 gets h_2 and is deleted from further rounds and preference orders.



Round 3 of TTC

$P_4 \longrightarrow h_5$, $P_5 \longrightarrow h_6$, $P_4 \longrightarrow h_4$ and we have completed the allocation.

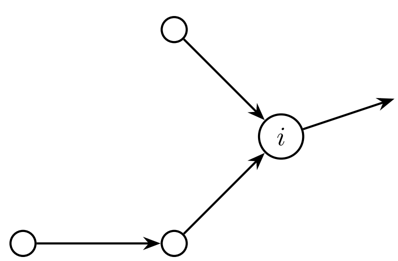
	P_1	P_2	P_3	P_4	P_5	P_6
$a^{(0)}$	h_1	h_2	h_3	h_4	h_5	h_6
a^{TTC}	h_3	h_2	h_1	h_5	h_6	h_4

2.1.3 Algorithmic form

- Initial endowment $a^{(0)}$
- $M_{remain} = M, N_{remain} = N$
- Repeat until $N_{remain} = \phi$
  1. Construct directed graph such that every agent points to its most favorite remaining house
  2. Find cycle in this graph and allocate houses to agents that are part of this cycle
  3. Remove the agents, houses, update M_{remain}, N_{remain}

2.2 Strategy-proofness and Efficiency of TTC

Take an arbitrary agent i, who gets a house in the round k of the TTC mechanism.



So in rounds $1, 2, \dots, k-1$, there was no closed loop with i in it. Now, for any round r, let $\Pi_{i,r} = \{$set of agents having a directed path to i in round r $\}$.

Observation: After each round, the set of agents having a directed path to i will be greater than or equal to that in the last round. In other words, $|\Pi_{i,r}| \geq |\Pi_{i,r-1}|$.

Theorem 2.1 *TTC with fixed endowment is* ***strategy proof****.*

Proof: Suppose, by truthful reporting P_i, i gets a house in round k. Let the false order that guarantees i a better allocation be $P_i^{'}$.

Case-1: i gets a house after round k Earlier, he pointed to the house of higher preference in round k and got it. But in a round after k, he is pointing to a house of lower preference. Therefore, he gets a house of strictly less priority.

Case-2: i gets a house before round k

If he gets anything before round k, say in round r, he is pointing towards a house in $\Pi_{i,r}$. The only way i can get assigned a house in round r is if i points to some house owned by an agent in $\Pi_{i,r}$(as other agents are not changing their actions, therefore i must create a cycle). Suppose $i \longrightarrow i^{'}, i^{'} \in \Pi_{i,r}$. Consider the path

$$i^{'} \longrightarrow i_1 \longrightarrow i_2 \longrightarrow \dots i_l \longrightarrow i$$

Each of these agents are pointing to their most favourite houses. If i does not point to $i^{'}$ in round r, then in round k, all these options will still be available. These houses won't get allocated anyway till k^{th} round. The fact that agent i's true preference in round k is none of these implies that the house i gets in round k is better than all of these. So, agent i gets an inferior house if it points to $i^{'}$.

Thus, any deviation from the true preference order leads to agent i recieving a house that is weakly dominated by the original allocation. Hence, TTC is strategy proof. ∎

Theorem 2.2 *TTC with fixed endowment is* ***Pareto efficient****.*

Proof: Suppose allocation a is done by TTC and $a^{'}$ is some other allocation such that $a^{'} \neq a$ and $a_i^{'} P_i a_i$ or $a_i^{'} = a_i \; \forall i$, i.e., every agent gets a better house or the same house than TTC in $a^{'}$.

Suppose, i is the agent who gets the first house that is different from TTC under $a^{'}$. Therefore $a^{'} \neq a_i$, and by assumption $a_i^{'} P_i a_i$. Since the houses allocated before i got its allocation under TTC are exactly the same, $a_i^{'}$ was available when a_i was assigned to i. But that's impossible under TTC. It always gives the most preferred house in that round. This leads to contradiction.

Hence, TTC is Pareto Efficient.

∎

Based on these 2 properties, TTC seems equivalent to Serial Dictatorship. **Stability** is a property that distinguishes between TTC and Serial dictatorship.

2.3 Stability of TTC

Consider the following allocation:

	P_1	P_2	P_3	P_4
$a^{(0)}$	h_1	h_3	h_2	h_4
a	h_1	h_2	h_3	h_4

Preference order for P_3 and P_4:

P_3	P_4
h_4	h_2
$\vdots$	$\vdots$
h_3	h_4

P_3 and P_4 have each other's top houses and they can just exchange house and break off from this allocation.

Group Deviation: Agents {3,4} can stay out of this allocation algorithm and atleast one of them gets strictly better off and both of them are weakly better off. This is called Group Deviation.

Some helpful definitions

- $\mathbf{a^{(0)}}$: Matching reflecting the initial endowment
- $\mathbf{a^S}$: Matching (allocation) of agents in $S \subseteq N$ over houses owned by agents in S.
- **Blocking coalition:** A coalition $S \subseteq N, S \neq \phi$ blocks a matching $'a'$ at a preference profile P if $\exists$ a matching a^S such that

$$\text{either} \quad a_i^S P_i a_i \quad \text{or} \quad a_i^S = a_i \quad \forall i \in S$$
$$\text{and}$$
$$\exists j \in S \quad a_j^S P_j a_j$$

- **Core:** An allocation/matching is in the "core" of a profile P if no coalition can block it at P
- **Stability:** SCF f is *stable* if $f(P)$ is in the core of P, $\forall P$

Stability ensures no blocking coalition for every size of coalitions (trivially including the grand coalition) which implies Pareto efficiency.

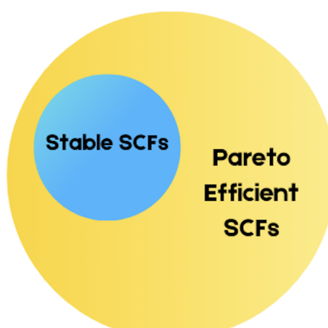


Figure 2.1: Stable SCFs form a subset of Pareto efficient SCFs

Example of a Pareto Efficient allocation that is not stable:

Initial allocation: $P_1 \longrightarrow h_1, P_2 \longrightarrow h_2, P_3 \longrightarrow h_3$

	P_1	P_2	P_3	
$a^{(0)}$	h_1	h_3	h_2	via serial dictatorship
a	h_1	h_1	h_3	

Preference order:

P_1	P_2	P_3
h_2	h_2	h_2
h_1	h_1	h_1
h_3	h_3	h_3

P_3 will block the allocation from serial dictatorship because he already has his best preference.

Theorem 2.3 *TTC is a stable allocation mechanism.*

Proof: Using proof by contradiction, suppose TTC is not stable. Therefore $\exists$ a blocking coalition at some profile P, lets call it S. Hence $\exists$ some allocation a^S such that

$$a_i^S P_i a_i \quad \text{or} \quad a_i^S = a_i \quad \forall i \in S \quad \text{and} \quad \exists j \in S \text{ s.t. } a_j^S P_j a_j$$

$T \subseteq S$ be the set of agents that get a strictly better house via the blocking coalition,

$$T = \{j \in S : a_j^S P_j a_j\} \neq \phi$$

1. Consider the agents from S who got allocated in R_1 of TTC $\longrightarrow S_1$. These agents cannot be a member of the set T because all those who got allocated are getting their top priorities in the first round.
2. For round 2, let $S_2 \subseteq S$ be the set of agents that get allocated. Any member of T cannot be in S_2 because that would imply that a^S for that agent was a house allocated in the previous round. But, since the agents who were allocated in the first round got their highest priority, any different allocation would be strictly dominated by their current allocation, so they would not want to deviate. So S_2 get their next top preferences and cannot improve it, thus $S_2 \cap T = \phi$
3. Applying induction on this argument, we can conclude that the set T is indeed empty, leading to contradiction. Hence, TTC is stable.

■

Theorem 2.4 *There is a unique core matching for every strict preference profile and initial endorsement and TTC find it.*

Proof: Suppose it is not unique, let there be an alternate core allotment $a^{'}$, with $a^{'} \neq a^{TTC}$

Follow the steps of TTC, let k^{th} round be the first round where $\exists i \; s.t. \; a_i^{'} \neq a_i^{TTC}$. By TTC mechanism, a_i is the most preferred remaining house for i, $a_i^{TTC} P_i a_i^{'}$.

Let S be the set of all agents that are allocated a house in the k^{th} round. Each agent will either be allotted the same house in alternative allotment($a_j^{'} = a_j^{TTC}$) or will get a strictly worse house in the alternative allotment($a_i^{TTC} P_i a_i^{'}$). These agents can form a blocking coalition, which contradicts our assumption.

Therefore, there is a unique core matching for every strict preference profile and initial endorsement. ■