

Algorithmic Game Theory- Top Trading Cycle Procedure

Michael Levet

June 5, 2015

1 Introduction

The House Allocation Game is a specific type of cooperative game which seeks to match actors with an equal number of houses. This solution concept associated with cooperative games is known as the core, which is a generalization of the Nash equilibrium. The best allocations in the core constitute what is known as the strong core. As it turns out, when preferences are strict, the strong core for this game contains a unique element.

This tutorial will examine the House Allocation Game. In particular, I will discuss the Top-Trading Cycle Algorithm to solve for the strong core allocation, as well as some results on this algorithm.

2 House Allocation Game

Prior to exploring the Top-Trading Cycle Algorithm, it is important to become familiar with the structure of the Housing Allocation Game. This section serves to provide that introduction.

The Housing Allocation Game has n players. Each player has an initial endowment of one house. We denote player i 's initial endowment as $\omega^i = h_i$. Each player also has a strict preference relation \succ^i over the houses, which he seeks to maximize. That is, every player wants his best house.

Consider the following example with three players, having the following preference profiles:

- p_1 : h_3, h_2, h_1
- p_2 : h_1, h_2, h_3
- p_3 : h_2, h_3, h_1

Recall that $\omega^1 = h_1, \omega^2 = h_2$, and $\omega^3 = h_3$. So consider the initial allocation $\omega = (h_1, h_2, h_3)$. We ask ourselves if there is a way for some of the players to swap their houses without leaving anyone worse off. Suppose p_1 and p_3 swap houses. Then the allocation is (h_3, h_2, h_1) , which leaves p_1 with his best choice and p_3 with his worst choice. However, since $\omega^3 = h_3 \succ^3 h_1$, this exchange isn't reasonable. The player p_3 can easily and realistically refuse to trade.

Suppose instead that p_1 and p_2 swap houses. This yields the allocation (h_2, h_1, h_3) , leaving p_1 with $h_2 \succ^1 \omega^1 = h_1$, p_2 with $h_1 \succ^2 \omega^2 = h_2$, and p_3 's allocation unchanged from ω^3 . So it is both beneficial and realistic for p_1 and p_2 to trade, and p_3 can't argue against this trade as it doesn't impact him.

The allocation (h_2, h_1, h_3) is still not the best allocation though. Players p_1 and p_3 can swap their houses under this new allocation to improve further, yielding the allocation (h_3, h_1, h_2) . Observe that each player has his most preferred house under this allocation.

The question now becomes how to allocate the houses amongst players such that each player will be reasonably happy. In order to answer this question, it is necessary to first understand precisely what it means for each player to be reasonably happy with an allocation. Recall from basic notions of Game Theory that the solution concept is the Nash equilibrium, which states that no individual can unilaterally deviate and improve his outcome. This notion of unilateral deviation does not capture the cooperative behaviors of the players, though.

The solution concept in cooperative games known as the core generalizes the notion of the Nash equilibrium. Intuitively, an allocation is in the core if no coalition that can pool their resources to obtain a better allocation for themselves. Let's now formalize these notions some.

Let $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$ be allocations. The allocation y is said to dominate x if there exists a coalition $S \subset \{1, 2, \dots, n\}$ such that $y_i \succeq^i x_i$ for all $i \in S$; and for at least one $j \in S$, $y_j \succ^j x_j$. That is, y weakly improves upon x for all members of the coalition S and strictly improves at least one coalition member's outcome. The coalition S is said to be a blocking coalition. An allocation x is said to be in the core if there exists no blocking coalition for x .

Recall the example above. There are only two core allocations: (h_3, h_1, h_2) and (h_2, h_1, h_3) . It was shown above that neither of these allocations have blocking coalitions. Recall as well that p_1 and p_3 prefer (h_3, h_1, h_2) while p_2 is indifferent between the two allocations. The question now arises how to categorize these two allocations. To this end, we utilize the strong core. Let's begin with some terminology.

First, we define the weak core. The weak core contains allocations such that no coalition $S \subset \{1, \dots, n\}$ exists that can construct a feasible allocation y satisfying $y_i \succ^i x_i$ for all $i \in S$. That is, the weak core contains allocations where no coalition can strictly improve each member's outcome.

The strong core contains allocations such that no coalition can weakly improve its members' outcomes. That is, there does not exist a coalition $S \subset \{1, \dots, n\}$ which can construct an allocation y satisfying $y_i \succeq^i x_i$ for all $i \in S$; and for at least one $j \in S$, $y_j \succ^j x_j$. Observe that the strong core is a subset of the weak core. Note as well that allocations in the strong core are overall better than those strictly in the weak core. Intuitively, more players are better off holistically under strong core allocations than in weak core allocations. Formally, a weak core allocation requires that no coalition can strictly better all of its members. In contrast, a strong core allocation simply requires that no coalition can strictly better an individual while leaving the other coalition members no worse for the wear.

Note: The definition of the core given above contains both the strong core and the weak core. Many authors adopt the convention of defining the core as I defined the weak core. The definition I provided of the strong core is consistent with more mainstream textbooks. I deviated in defining the core for the sake of clarity.

So now we return to the two core allocations (h_3, h_1, h_2) and (h_2, h_1, h_3) . Observe that the grand coalition $S = \{1, 2, 3\}$ weakly improves upon (h_2, h_1, h_3) by allowing p_1 and p_3 to exchange the houses h_2 and h_3 , strictly improving their outcomes. However, the grand coalition does not strictly improve p_2 's outcome, as p_2 already has his top choice- h_1 . Thus, (h_2, h_1, h_3) is in the core, but not the strong core. However, (h_3, h_1, h_2) is in the strong core, as each player has his top choice. Thus, no coalition can strictly improve at least one player's outcome.

3 Top-Trading Cycle Procedure

This section will introduce the Top-Trading Cycle Procedure, which returns the unique strong core allocation in the Housing Allocation Game. The algorithm starts by having each player reference the owner of his preferred house, where the house owners are determined by the initial allocations. For example, player p_1 owns h_1 . After each player references the owner of his preferred house, the algorithm seeks out a cycle and assigns each player along the cycle his preference. The algorithm then repeats this process with the remaining available players and unassigned houses until no player remains.

Let's work through the example from above. In the first iteration, p_1 references p_3 as p_1 's top choice is h_3 . Next, p_2 references p_1 . Finally, p_3 references p_2 . So we have a cycle $p_1 \rightarrow p_3 \rightarrow p_2 \rightarrow p_1$. So we assign h_3 to p_1 ; h_2 to p_3 ; and h_1 to p_2 , which is the strong core allocation for this instance of the Housing Allocation Game.

Now let's consider a more complicated example. Suppose there are five players p_1, \dots, p_5 with the following preference profiles:

- $p_1 : h_4, h_3, h_2, h_1, h_5$

- $p_2 : h_4, h_1, h_2, h_3, h_5$
- $p_3 : h_1, h_4, h_3, h_2, h_5$
- $p_4 : h_3, h_2, h_1, h_4, h_5$
- $p_5 : h_1, h_5, h_2, h_4, h_3$

At a given iteration, each player points to his top preference. We then remove cycles and repeat the process:

- **Iteration 1:** We have $p_1 \rightarrow p_3, p_2 \rightarrow p_4, p_3 \rightarrow p_1, p_4 \rightarrow p_3$, and $p_5 \rightarrow h_1$. Observe that there is the cycle $p_1 \rightarrow p_3 \rightarrow p_1$. So we allocate h_3 to p_1 ; and h_1 to p_3 .
- **Iteration 2:** We have $p_2 \rightarrow p_4$. Next, $p_4 \rightarrow p_2$, as h_3 has already been allocated. Finally, $p_5 \rightarrow p_5$, as h_1 has already been allocated. Observe the cycle $p_4 \rightarrow p_2 \rightarrow p_4$, so we allocate h_4 to p_2 and $h_2 \rightarrow p_4$.
- **Iteration 3:** Finally, we have p_5 by itself. So we allocate h_5 to p_5 .

Thus, the strong core allocation is $(h_3, h_4, h_1, h_2, h_5)$.

Now let's examine an implementation of the Top Trading Cycle Procedure. First, there is the Actor class, which models a player in the House Allocation Game. Each Actor object has a preference profile, a matched house, and a name. The Actor is assigned an initial house, as well.

SEE TUTORIAL FOR Actor.java code

Next, consider the AllocationMechanism class, which implements the Top Trading Cycle Procedure given a List of Actors. The AllocationMechanism class starts by constructing a HashSet corresponding to the unallocated houses. This adds a measure of efficiency during the house allocation process, as it can be checked in constant time if an Actor's most preferred house is available. The allocateHouses() method begins the Top Trading Cycle Procedure. It defers to the methods determineTopPicks() to obtain each player's most preferred available house. The methods findCycleRoot() and constructCycle() are then used to find a cycle amongst the players' preferences. The allocateHouses() method then allocates the appropriate houses to the corresponding players' in the cycle, removing both the allocated houses and satiated players from further consideration.

Note that the determineTopPicks() method performs an optimization. If it finds a player who prefers his initial allocation, it removes the player from further consideration, cutting down on unnecessary iterations.

SEE TUTORIAL FOR AllocationMechanism.java code

The TTCP class is the driver class for this program, demonstrating the five player example from above.

SEE TUTORIAL FOR TTCP.java code

Now that the Top-Trading Cycle Procedure has been introduced, the next step is to prove its correctness.

Theorem 1: The Top-Trading Cycle Procedure terminates; and upon termination, it returns the unique strong core allocation.

In order to prove termination, the following lemma will be introduced.

Lemma 1.1: Let $n \in \mathbb{N}$ and let $(p_i, h_i, \succ^i)_{i=1}^n$ be an instance of the Housing Allocation Game. Under the first iteration of the Top Trading Cycle Procedure, there exists a cycle amongst the players' preferences.

Proof: Suppose to the contrary that no cycle exists amongst the player preferences. Let p_i be a player. Since each player has a strict preference relation and no cycle exists amongst the player preferences, there exists a second player p_j whose house p_i most prefers. It follows that there exists a directed Hamiltonian path amongst

the players' most preferred preferences. By the pigeonhole principle, the last player on the directed Hamiltonian path must point to another player on (including himself) by assumption of having preferences over the houses. This constructs a cycle, a contradiction. QED.

Claim 1.1: The Top-Trading Cycle Procedure terminates.

Proof: By Lemma 1.1, on the first iteration of the Top-Trading Cycle Procedure, there exists a cycle amongst the players' preferences. Floyd's algorithm successfully finds one such cycle. The Top-Trading Cycle Procedure then assigns all players along this cycle their preferred houses and removes the players. As the cycle corresponds to housing preferences, only the houses initially owned by players on the cycle are allocated. It follows that the remaining players and houses constitute a new instance of the House Allocation Game. So we iterate on the above argument. As there are a finite number of players in the original instance of the House Allocation Game, the Top-Trading Cycle Procedure must terminate. QED.

Claim 1.2: The allocation produced by the Top-Trading Cycle Procedure is in the strong core.

Proof: It suffices to show that no weakly improving coalition exists. It will be shown by induction on $n \in \mathbb{N}$, the number of iterations of the algorithm, that no player removed on iteration n can participate in a weakly improving coalition. Consider the base case of $n = 1$. Then every player removed obtains his best choice; and thus, has no incentive to participate in a weakly improving coalition. Furthermore, as preferences are strict, top preferences are unique. If a player not matched on iteration 1 has a top preference outside of the players matched on this iteration, then this player has no incentive to block the Top-Trading Cycle allocation at this iteration. If instead a player's top preference was within the cycle removed on iteration one, then this player j can only strictly improve his preference by obtaining the house of a player i in this cycle. This in turn displaces a player k in the cycle, who would then form a blocking coalition with his cycle to block the allocation induced by j . So at iteration 1, no coalition can weakly improve upon the Top-Trading Cycle allocation.

Suppose the theorem holds true up to $k > 1$ iterations of the Top-Trading Cycle Procedure. The $k + 1$ case will now be shown. By the inductive hypothesis, no player matched in the first k iterations can participate in a weakly improving coalition. If all players have been matched, then we are done. Otherwise, let player x be an unmatched player after $k + 1$ iterations of the Top-Trading Cycle Procedure. If player x 's top choice house is not amongst the matched players' initial allocations, then player x has no incentive to block the Top-Trading Cycle Procedure allocation. Suppose instead that player x 's top preferred house is amongst the matched houses. Then obtaining this house displaces an already matched player, who could strictly improve upon the housing allocation induced by player x by forming a blocking coalition with his cycle. It follows by induction that no weakly improving coalition exists for the Top-Trading Cycle Procedure allocation. QED.

Claim 1.1 and Claim 1.2 together imply Theorem 1.

Theorem 2: The Top-Trading Cycle Procedure executes in $O(n^2)$ time.

Proof: In the worst case, there are $\lfloor n/2 \rfloor$ cycles, each containing two players. Thus, there are $O(n)$ cycles to detect. The adaptation of Floyd's cycle detection algorithm takes $O(n)$ time, and matching the players in the cycle takes $O(n)$ time. Thus, $T(n) = O(n) \cdot (O(n) + O(n)) = O(n^2)$. QED.

4 Results on Top-Trading Cycle Procedure

This section will introduce additional results on the Top-Trading Cycle algorithm. The goal is to examine the economic issues of the Top-Trading Cycle Procedure rather than the computer science issues. It will first be shown that the strong core is unique. The second topic of interest deals with evaluating how well the Top Trading Cycle Procedure performs as a mechanism. We analyze the incentive for players to honestly reveal

their preferences, as well as the existence (or lack thereof) of more economically efficient mechanisms to solve the House Allocation Game.

Let's begin with the first result- the uniqueness of the strong core.

Theorem 3: In the Housing Allocation Game, the strong core has a single allocation.

Proof: Suppose to the contrary that there exist at least two allocations in the strong core. Let x be the strong core allocation returned by the Top Trading Cycle Procedure, and let y be a second distinct strong core allocation. Let $S = \{i : x_i \neq y_i\}$. As $x \neq y$, $|S| \geq 2$. If there exists a player $i \in S$ such that $x_i \succ^i y_i$, then player i will form a weakly improving coalition with the players from his cycle under the Top-Trading Cycle Procedure. This implies that all players in S must strictly improve upon their preferences. Let $i \in S$ such that $y_i \succ^i x_i$. Let player i be matched in iteration k of the Top-Trading Cycle Procedure. Since $y_i \succ^i x_i$, it follows that player i preferred the house of some other player j to his own initial allocation. By Claim 1.2, this player j must have been matched prior to iteration k , and y displaces a player matched in the same iteration as player j . Thus, the players matched in the same iteration as player j all form a weakly improving coalition. Thus, $S \neq \emptyset$ implies the existence of a weakly improving coalition blocking y . It follows that x is the unique strong core allocation. QED.

The next result shows that the Top-Trading Cycle Procedure is strategy proof. That is, it incentivizes all players to honestly report their preferences. In fact, there is a stronger result proven by Jinpeng Ma in 1994: the Top-Trading Cycle Procedure is the unique mechanism that guarantees individual rationality, Pareto optimality, and strategy proofness in the house allocation game.

Theorem 4: The Top-Trading Cycle Procedure is strategy proof.

Proof: Let p be a player, and let k be the iteration at which p was assigned a house. If p was assigned his top choice, then p has no incentive to dishonestly report his preferences. Suppose instead player p does not receive his top choice. Then p can only improve by obtaining a house allocated in an earlier iteration. Misrepresenting his preferences will not create a new directed cycle, so p will not be able to obtain a house allocated in an earlier iteration. It follows that it is a dominant strategy for p to report his preferences honestly. As the choice of p was arbitrary, the result holds for all players. QED.

We next seek to prove that the Top-Trading Cycle Procedure is the unique mechanism that guarantees individual rationality, Pareto optimality, and strategy proofness in the house allocation game. Let's first introduce the definitions of individual rationality and Pareto optimality. Individual rationality is quite intuitive. A decision is individually rational if no player is worse off than under his initial allocation. In the House Allocation Game, individual rationality implies that a player will not participate in a coalition unless his outcome is at least as good as keeping his initial house.

Pareto optimality is quite intuitive as well. An allocation x is Pareto optimal if no other allocation y can weakly improve upon x . Formally, y weakly improves upon x if $y_i \succeq^i x_i$ for all players i ; and $y_j \succ^j x_j$ for at least one player j . Observe that the Core is a subset of the Pareto optimal allocations.

The result that the Top-Trading Cycle Procedure is the unique mechanism that guarantees individual rationality, Pareto optimality, and strategy proofness in the house allocation game quite important, as it sets this mechanism apart from many alternative. Consider the Serial Dictatorship mechanism, which assigns a given player his best available choice. This mechanism is strategy proof and Pareto optimal. However, it is not individually rational. Certain orderings of the players may leave individuals worse off than their initial allocations. The Top-Trading Cycle Procedure works quite similarly to the Serial Dictatorship mechanism. So the fact that the Top-Trading Cycle Procedure is not only individually rational, Pareto efficient, and strategy proof; but also the unique mechanism satisfying these properties is an impressive result.

Theorem 5: A house allocation mechanism is individually rational, Pareto optimal, and strategy proof if and only if the mechanism is the Top-Trading Cycle Procedure.

Claim 5.1: The Top-Trading Cycle Procedure is individually rational, Pareto optimal, and strategy proof.

Proof: Let N represent the set of players, with $\succ := (\succ^i)_{i=1}^{|N|}$ the corresponding preferences. Suppose that the mechanism used is the Top-Trading Cycle Procedure. The allocation yielded by this mechanism is in the strong core, which implies individual rationality and Pareto optimality. Theorem 4 provides the result that the Top-Trading Cycle Procedure is strategy proof. QED.

Showing that the Top-Trading Cycle Procedure is the unique mechanism satisfying individual rationality, Pareto optimality, and strategy proofness requires more work. The goal is to start with an arbitrary mechanism, which is assumed to satisfy individual rationality, Pareto optimality, and strategy proofness. We then seek to contradict strategy proofness by providing an instance where a player can misrepresent his preferences and improve his payoff. Let's start with some Lemmas, which will be helpful later.

Lemma 5.1: Let x be the strong core allocation and let y be both individually rational and Pareto optimal, with $x \neq y$. Then there exists a player i such that $x_i \succ^i y_i \succ^i \omega^i$.

Proof: Suppose to the contrary that for all players $i \in N$, $x_i \succ^i y_i \succ^i \omega^i$ is false. As both x, y are Pareto optimal and preferences are strict, there exist players i, j such that $x_i \succ^i y_i$ and $y_j \succ^j x_j$. For every player k that prefers $y_k \succ^k x_k$, it follows that $y_k = \omega^k$. Let $S = \{i \in N : x_i \sim^i y_i \text{ or } y_i \succ^i x_i\}$ be a coalition. As y is individually rational for all players, S weakly improves upon x , contradicting the assumption that x is the strong core allocation. QED.

Let $S_P = \{i \in N : \text{there exists } j \in N \text{ s.t. } \phi_i(N, \succ) \succ^i h_j \succ^i \omega^i\}$. Now we define a preference relation \succ' of misrepresented preferences. If player $i \notin S_P$, then player i does not misrepresent his preferences. Now suppose $i \in S_P$. Let k be the index in \succ^i for the house $TTC(N, \succ)(i)$. Then $(\succ'_i)_{i=1}^k = (\succ_i)_{i=1}^k$. Player i then reports $\succ'_{k+1} = \omega^i$. Finally, player i reports the remaining houses in preference order of \succ^i following ω^i .

Remark 1: Observe that by construction of \succ^i , $TTC(N, \succ) = TTC(N, \succ') = TTC(N, \succ_S, \succ'_{-S})$ for any $S \subset N$. That is, the strong-core allocation is returned by the Top-Trading Cycle Procedure regardless of which subset of players reports preferences dishonestly according to \succ' .

It will now be shown that any mechanism ϕ satisfying individual rationality, Pareto optimality, and strategy proofness has the property that $\phi(N, \succ') = TTC(N, \succ')$.

Lemma 5.2: Let ϕ be a mechanism satisfying individual rationality, Pareto optimality, and strategy proofness. Then $\phi(N, \succ') = TTC(N, \succ')$.

Proof: Suppose to the contrary that $\phi(N, \succ') \neq TTC(N, \succ')$. By Lemma 5.1, let player i such that:

$$TTC(N, \succ')(i) \succ'_i \phi(N, \succ')(i) \succ'_i \omega^i.$$

By construction of \succ' and the fact that:

$$TTC(N, \succ) = TTC(N, \succ') = TTC(N, \succ_S, \succ'_{-S})$$

We have that for every player $j \in N$, ω^j immediately follows: $TTC(N, \succ')(j)$ in \succ'_j or $TTC(N, \succ')(j) = \omega^j$.

Either case contradicts:

$$TTC(N, \succ')(i) \succ'_i \phi(N, \succ')(i) \succ'_i \omega^i$$

QED.

Finally, we get to the backbone for the converse of the main result:

Lemma 5.3: Let ϕ be a mechanism satisfying individual rationality, Pareto optimality, and strategy proofness. Then $\phi(N, \succ'_{-S}, \succ_S) = TTC(N, \succ'_{-S}, \succ_S)$ for every $S \subset N$.

Proof: By Lemma 5.2, it suffices only to consider players in S_P , as players in $N \setminus S_P$ report their preferences in \succ' as defined in \succ . The proof is by induction on $|S|$, where $S \subset S_P$. Consider the base case of $|S| = 0$. Then all players report \succ . By Lemma 5.2, we have $\phi(N, \succ) = TTC(N, \succ)$ as desired.

Suppose that this lemma holds up to an arbitrary $|S| = k$. Consider the case where $|S| = k + 1$. Let $Q = (\succ'_{-S}, \succ_S)$. Suppose that $TTC(N, Q) \neq \phi(N, Q)$. By Lemma 5.1, there exists a player i such that:

$$TTC(N, Q)(i) \succ_Q^i \phi(N, Q)(i) \succ_Q^i \omega^i$$

Where \succ_Q^i denotes player i 's preference relation under Q . Suppose first $i \in N \setminus S$. From Remark 1, recall that $TTC(N, Q)(i) = TTC(N, \succ)(i)$. From the proof of Lemma 5.2, this contradicts the construction of \succ' , as either ω^i follows $TTC(N, \succ)(i)$ or $TTC(N, \succ)(i) = \omega^i$.

Suppose instead $i \in S$. From Remark 1, we have that $TTC(N, Q)(i) = TTC(N, Q_{-i}, \succ'_i)(i)$. By the inductive hypothesis, $TTC(N, Q_{-i}, \succ'_i)(i) = \phi(N, Q_{-i}, \succ'_i)(i)$.

Recall that under Q , player i was honestly reporting his preferences. It follows that $\phi(N, Q_{-i}, \succ'_i)(i) \succ^i \phi(N, Q)(i)$, where \succ^i is player i 's actual preference relation. Thus, under ϕ , player i benefits by dishonestly reporting his preferences, contradicting the assumption that ϕ is strategy proof. QED.

Claim 5.2: Let ϕ be a mechanism to solve the House Allocation Game that is individually rational, Pareto optimal, and strategy proof. Then $\phi(N, \succ) = TTC(N, \succ)$. That is, the Top-Trading Cycle Procedure is the unique mechanism satisfying individual rationality, Pareto optimality, and strategy proofness.

Proof: We apply Lemma 5.3, setting $S = N$. QED.

Claims 5.1 and 5.2 together imply Theorem 5.