MATCHING WITH RENEGOTIATION

AKIHIKO MATSUI AND MEGUMI MURAKAMI

Abstract. The present paper studies the deferred acceptance algorithm (DA) with renegotiation. After DA assigns objects to players, the players may renegotiate. The final allocation is prescribed by a renegotiation mapping, of which value is a strong core given an allocation induced by DA. We say that DA is *strategy-proof with renegotiation* if for any true preference profile and any (potentially false) report, an optimal strategy for each player is to submit the true preference. We then have the first theorem: DA is not strategy-proof with renegotiation for *any* priority structure. The priority structure is said to be unreversed if there are two players who have higher priority than the others at any object, and the other players are aligned in the same order across objects. Then, the second theorem states that the DA rule is implemented in Nash equilibrium with renegotiation, or equivalently, any Nash equilibrium of DA with renegotiation induces the same allocation as the DA rule, if and only if the priority structure is unreversed.

Keywords: deferred acceptance algorithm (DA), indivisible object, renegotiation mapping, strong core, strategy-proofness with renegotiation, Nash implementation with renegotiation, unreversed priority

JEL Classification Numbers: C78, D47

1. Introduction

Deferred acceptance algorithm (DA) has been studied since the seminal paper by Gale and Shapley (1962). A class of problems investigated therein includes college admissions and office assignments. In these problems, the centralized mechanism assigns objects to players based on their preferences over objects and priorities for the objects over the players. Two notable properties of DA are stability and strategy-proofness (Gale & Shapley, 1962; Dubins & Freedman, 1981; Roth, 1982). Because of these properties, DA plays a significant role both theoretically and practically.

Date: March 2, 2020.

Incomplete.

We are grateful to the participants of seminars at Keio University, Kyoto University, Kobe University, National University of Singapore, Northwestern University, Okayama University, Waseda University, Econometric Society Summer School, and UT Summer School in Economics, as well as In-Koo Cho, Ryota Iijima, Mamoru Kaneko, Daiki Kishishita, Fuhito Kojima, Shunya Noda, Wojciech Olszewski, Ariel Rubinstein, Marciano Siniscalchi, Bruno Strulovici, and Yuki Takagi for useful conversations and comments. Financial support from the Grants-in-Aid for Scientific Research for the Promotion of Science is gratefully acknowledged. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Japan Society for the Promotion of Science.

The studies on DA typically assume, often implicitly, that the players can commit to the object allocation of the mechanism, *i.e.*, renegotiation is prohibited[1]. While this assumption is not unreasonable in certain settings such as college admissions, it is not clear why it should *necessarily* be the benchmark assumption in some other cases such as office assignments.

The present paper explicitly incorporates renegotiation into DA and analyzes its properties. In the bold stroke, the present analysis is described as follows. To begin with, it focuses on the deferred acceptance algorithm (DA) of Gale and Shapley (1962)[2]. There are finitely many indivisible objects and players. Each object a is endowed with a priority relation $\succ_a$ over the players, where $i \succ_a j$ implies that player i has higher priority than player j at a. On the other hand, each player i has a preference P_i over the objects, where aP_ib implies that player i strictly prefers a to b. There is no tie for both priority and preference relations. The objects are assigned through DA based on priority as well as the preference profile submitted by the players. Each player obtains one and only one object. Given a priority structure $\succ$ and a (potentially false) report P, the DA allocation for P is denoted by $DA_{\succ}(P)$.

After the assignment through DA, the players cannot commit to the allocation, *i.e.*, they may renegotiate for a better allocation. This renegotiation is characterized by a renegotiation mapping r, which is similar in spirit to, *albeit* different from, the one defined in the literature on mechanism with renegotiation (see, *e.g.*, the survey by Maskin & Sjöström, 2002). The outcome after renegotiation is assumed to be a strong core given the (interim) allocation of DA as an endowment. Their function assumes that the outcome is Pareto efficient and individually rational, but not necessarily a strong core. This renegotiation mapping is well-defined due to Shapley and Scarf (1974), which shows the non-emptiness of core, and Roth and Postlewaite (1977), which shows the unique existence of strong core[3]. Given a true preference profile P and a (potentially false) report P', the final allocation after renegotiation is denoted by $r(DA_{\succ}(P'); P)$. Note that the final allocation depends upon DA and renegotiation, where the former is based on the report P', while the latter is on the true preference profile P.

Then, given $\succ$, $DA_{\succ}$ is said to be *strategy-proof with renegotiation* if for all P and all P', for any player i, and any report P''_i, we have

$$r_i(DA_{\succ}(P_i, P'_{-i}); P)\ R_i\ r_i(DA_{\succ}(P''_i, P'_{-i}); P).$$

In other words, the true preference P_i of player i has to be a best response to any (true) preference P_{-i} and any (potentially false) report P'_{-i} of the other players[4].

The first theorem states that $DA_{\succ}$ is *not* strategy-proof with renegotiation for any priority structure $\succ$.

[1]In the theory of two-sided matching markets, Hatfield and Kojima (2010) considers renegotiation between doctors and hospitals. However, the notion of renegotiation therein is more related to stability rather than a formal concept in the literature of mechanism design with renegotiation (see, *e.g.*, Maskin & Sjöström, 2002).

[2]Game-theoretic formulation of DA is taken from Sotomayor (2008).

[3]Kaneko (1982) proves the non-emptiness of core under non-transferable utility. Wako (1984) shows that a strong core is inside the set of competitive equilibria and demonstrates the conditions under which a strong core exists.

[4]In the standard definition of strategy-proofness without renegotiation, $DA_{\succ}$ is said to be *strategy-proof* (without renegotiation) if for any P, for any player i, and any report P'_i, we have $DA_{\succ,i}(P)\ R_i\ DA_{\succ,i}(P'_i, P_{-i})$. Note that in this definition, it does not matter whether P_{-i} is the true preference profile of the others or not.

For the sake of illustration, consider the following example. There are three players, 1, 2, and 3, and three objects, x, y, and z. The priority structure is given by

$$2 \succ_x 1 \succ_x 3,$$
$$1 \succ_y 2 \succ_y 3,$$
$$1 \succ_z 2 \succ_z 3.$$

Consider a preference profile P given as follows:

$$xP_1zP_1y,$$
$$yP_2zP_2x,$$
$$xP_3yP_3z.$$

DA will give the allocation (x, y, z), where players 1, 2, and 3 obtain x, y, and z, respectively. Suppose now that 1 misreports his preference and puts y on the top of the list, while 2 puts x on the top. Then, DA induces (y, x, z). Players 1 and 2 then renegotiate to reach the final allocation of (x, y, z). In this situation, player 1 is strictly better off by the misreport compared to the truth-telling. Indeed, if player 1 follows P_1, then player 1 will obtain z in DA, and the final allocation will be (z, y, x). Similarly, player 2 is strictly better off by the misreport, too. In this sense, the true preference is not a weakly dominant report. Thus, DA is not strategy-proof.

In the above example, however, it is still verified that the true preference profile is a Nash equilibrium, and that any Nash equilibrium induces the same outcome. We consider implementation in Nash equilibrium with renegotiation. Given $\succ$, let $\overline{DA}_\succ$ be the social choice rule that is given by the mechanism $DA_\succ$ *without* renegotiation[5]. We then find an equivalence condition for the implementation of $\overline{DA}_\succ$. The priority structure is said to be *unreversed* if there are two players who have higher priority than the others at any object, and the other players are aligned in the same order across objects. The second theorem states that $\overline{DA}_\succ$ is implemented in Nash equilibrium with renegotiation, or equivalently, given an arbitrary P, any Nash equilibrium P' of DA with renegotiation satisfies $r(DA_\succ(P'); P) = \overline{DA}_\succ(P)$, if and only if the priority structure is unreversed.

In the second theorem, DA is used for both the social choice rule and the mechanism that implements it. In the literature of mechanism design with renegotiation (*e.g.*, Maskin & Sjöström, 2002), a social choice rule is said to be implementable in Nash equilibrium with a renegotiation mapping if there exists a mechanism in which a Nash equilibrium induces the allocation designated by the rule. There are three differences in the present analysis. First, we confine our attention on DA to derive a specific equivalence condition. Second, we require that every pure strategy Nash equilibrium should induce the allocation designated by the rule. Third, the renegotiation mapping of the present paper is uniquely determined so that our definition of implementation does not depend on the selection of the mapping.

The DA rule is implemented in the first example because the priority structure is unreversed. If a priority structure is reversed, we have the following example. Consider a priority structure that satisfies:

$$1 \succ_x 2 \succ_x 3,$$
$$1 \succ_y 3 \succ_y 2.$$

[5]Note that for any $\succ$, and any P, $\overline{DA}_\succ(P) = DA_\succ(P)$ holds. Although they are mathematically identical, we use different symbols for the social choice rule and the mechanism that implements it to make our definitions and statements clear.

Suppose further that the preference profile is given by

$$xP_1yP_1z,$$
$$yP_2xP_2z,$$
$$yP_3xP_3z.$$

Then, for the true preference profile, the DA allocation is (x, z, y), and there will be no further renegotiation. If, however, player 1 misreports by putting y on the top of the list, then players 2 and 3 are rejected at y, and they end up with x and z, respectively, *i.e.*, the outcome of DA becomes (y, x, z). Then there will be renegotiation between 1 and 2, who will exchange their objects, and the final outcome becomes (x, y, z), which is different from the DA allocation induced by the true preference profile.

The rest of the paper is organized as follows. Section 2 introduces notation and definitions, including the description of DA and the definition of the renegotiation mapping. Section 3 shows that for any priority structure, DA is *not* strategy-proof with renegotiation. Section 4 gives an equivalence condition under which DA is implemented in Nash equilibrium with renegotiation. Section 5 mentions two remarks, one on an alternative definition of a renegotiation mapping, and the other on priority structures. Section 6 concludes the paper. Some proofs are relegated to appendices.

2. Model

2.1. **Preliminaries.**

N is a finite set of players. O is a finite set of objects. Assume $|N| \geq 3$ and $|O| \geq |N|$. The objects are indivisible, and each player demands at most one unit of object.

An object allocation is $\mu \in O^N$ that satisfies $[i \neq j \Rightarrow \mu_i \neq \mu_j]$. Let A denote the set of allocations. Let R_i denote i's preference preorder ($i \in N$) over O. For all $a, b \in O$, we write aP_ib if aR_ib holds, and bR_ia does not. P_i represents player i's strict preference. We may write $P_i = (a_i^1, \ldots, a_i^{|O|})$ where $a^nP_ia^{n'}$ for $n < n'$. We write $P = (P_i)_{i\in N}$. We assume that there is no tie, *i.e.*, $[xR_iy \wedge yR_ix] \Rightarrow x = y$. Let $\mathcal{P}_i$ be the set of all strict preferences of player i. We write $\mathcal{P} = (\mathcal{P}_i)_{i\in N}$. In the sequel, we write $P_i \in \mathcal{P}_i$, implying that the associated preference preorder is R_i.

2.2. **The assignment through the deferred acceptance algorithm.**

We consider the deferred acceptance algorithm (DA).

The players simultaneously submit their preferences; $P_i \in \mathcal{P}_i$ for player $i \in N$. For every $a \in O$, $\succ_a$ is a strict total order at $a \in O$ over N, *i.e.*, it satisfies transitivity and asymmetry, and has no non-comparable pairs[6]. It defines the order of players' priority at object a, *i.e.*, $i \succ_a j$ means that i has higher priority than j at a. Let $\succ = (\succ_a)_{a\in O}$ be a priority structure.

Given a priority structure $\succ$ and a report P with $P_i = (a_i^1, \ldots, a_i^{|O|})$ $(i \in N)$, the DA allocation $DA_\succ(P)$ is determined by DA of Gale and Shapley (1962):

Step 1: Player $i \in N$ applies to a_i^1.
For each $a \in O$, if only one player chooses a, then the player is *temporarily* assigned to a. If there are multiple players choosing a, then the player with the highest priority at a is *temporarily* assigned to a and applies to a in the next step, and the others are rejected.

[6] A binary relation $\succ_a$ over N is said to have *no non-comparable pairs* if $i \neq j$ implies either $i \succ_a j$ or $j \succ_a i$.

Step t ($t > 1$)**:** The players who were rejected at Step $t-1$ apply to their next object in their respective lists. For each $a \in O$, if there is only one player who applies to a, then this player is *temporarily* assigned to a. If there are multiple players choosing a, then the player with the highest priority at a is *temporarily* assigned to a and applies to a in the next step, and the others are rejected.

Termination: Terminate the process when all the players are assigned to an object in O.

2.3. **Renegotiation.** We assume that players cannot commit to the outcome of DA, and that they may renegotiate and exchange their objects given the outcome of DA as an endowment. To characterize such a renegotiation process, we first define S-feasibility of allocation given a subset S of players.

Definition 2.1. *Given an outcome $\omega \in A$, and $S \subset N$, $\mu \in A$ is S-feasible with respect to (w.r.t.) ω if*

$\{\mu_i |\ i \in S\} = \{\omega_i |\ i \in S\}$.

An allocation μ is said to be feasible w.r.t. ω if it is N-feasible w.r.t. ω.

We introduce the notion of blocking to consider the endowment that is subject to renegotiation.

Definition 2.2. *Given $P \in \mathcal{P}$, $\omega \in A$, and a non-empty subset $S \subset N$, an allocation $\nu \in A$ blocks another allocation μ via S w.r.t. (P, ω) if*

- *ν is S-feasible w.r.t. ω,*
- *$\nu_i R_i \mu_i$ for all $i \in S$,*
- *$\nu_i P_i \mu_i$ for some $i \in S$.*

An allocation ν is said to block μ if ν blocks μ via S for some S. S is said to be a blocking coalition of μ if there exists ν that blocks μ via S.

We then give the definition of strong core, which is an allocation which has no blocking coalition.

Definition 2.3. *Given $P \in \mathcal{P}$, $\omega \in A$, an allocation μ is a* strong core *w.r.t. (P, ω) if*

- *μ is feasible w.r.t. ω,*
- *there is no blocking coalition of μ w.r.t. (P, ω).*

The outcome after renegotiation is assumed to be a strong core given the (interim) allocation of DA as an endowment profile. Whenever players may renegotiate, they exchange their endowment based on their true preference profile. Therefore, the following renegotiation mapping is defined as a function of the true preference profile as well as the endowment.

Definition 2.4. *A mapping $r : A \times \mathcal{P} \to A$ is a* renegotiation mapping *if for all $P \in \mathcal{P}$, and all $\omega \in A$, $r(\omega; P)$ is a strong core w.r.t. (P, ω).*

The next lemma restates a result in Roth and Postlewaite (1977).

Lemma 2.1. *(Roth & Postlewaite, 1977) There exists a unique renegotiation mapping. Thus, the renegotiation mapping is well-defined.*

3. Strategy-proofness of DA with renegotiation

This section considers the strategy-proofness of DA with renegotiation. It is shown that DA is not strategy-proof with renegotiation.

To begin with, we define the strategy-proofness of DA with renegotiation given a priority structure.

Definition 3.1. *Given priority structure $\succ$, $DA_{\succ}$ is strategy-proof with renegotiation if for all $P \in \mathcal{P}$, for all $i \in N$, for all P'_{-i}, and all P''_i,*

$$r_i(DA_{\succ}(P_i, P'_{-i}); P)\ R_i\ r_i(DA_{\succ}(P''_i, P'_{-i}); P).$$

A couple of remarks are in order. First, in strategy-proofness, what is required is that submitting the true preference is weakly dominant. This implies that for each i, P_i is at least as good as another report P''_i irrespective of the true preferences P_{-i} and (potentially false) reports P'_{-i} of the other players. In the standard definition of strategy-proofness without renegotiation, we do not have to distinguish P_{-i} and P'_{-i} because once P'_{-i} is submitted, the outcome does not depend upon P_{-i}. In the presence of renegotiation, however, even after P'_{-i} is submitted, P_{-i} still matters because there is a renegotiation stage where the players renegotiate based on their true preferences P_{-i}. Thus, the definition is a little more complicated than the standard definition of strategy-proofness without renegotiation. Second, in this definition, strategy-proofness is defined for each priority structure. This reflects the idea that the priority structure is part of the mechanism and common knowledge among the players.

We are now in a position to state the following theorem.

Theorem 3.1. *For any priority structure $\succ$, $DA_{\succ}$ is not strategy-proof with renegotiation.*

Proof. See Appendix A. □

Its proof roughly goes as follows. We take an arbitrary priority structure and construct a preference profile P such that some player, say, i, can be better off by submitting a profile $P''_i \neq P_i$ when the other players submit P'_{-i}. This construction is similar to, but more comprehensive than, the first example provided in the introduction.

4. Nash implementation with renegotiation

This section examines implementation in Nash equilibrium with renegotiation. Given $\succ$, let $\overline{DA}_{\succ}$ be the social choice rule that is implemented by the mechanism $DA_{\succ}$ without renegotiation. Note that for any $\succ$, and any P, $\overline{DA}_{\succ}(P) = DA_{\succ}(P)$ holds. Although they are mathematically identical, we use different symbols for the social choice rule and the mechanism that implements it to make our definitions and statements clear. We then have the following two definitions. First, the definition of Nash equilibrium with renegotiation is given.

Definition 4.1. *Given P, P' is a Nash equilibrium with renegotiation for P if for all $i \in N$, and all P''_i,*

$$r_i(DA_{\succ}(P'); P)\ R_i\ r_i(DA_{\succ}(P''_i, P'_{-i}); P)$$

Second, we have the definition of implementability in Nash equilibrium with renegotiation. We require that any Nash equilibrium with renegotiation induces the DA allocation for the true preference profile.

Definition 4.2. *Given $\succ$, $\overline{DA}_{\succ}$ is implemented in Nash equilibrium with renegotiation if for all P,*

- *a Nash equilibrium with renegotiation for P exists;*
- *if P' is a Nash equilibrium with renegotiation for P, then $r(DA_{\succ}(P'); P) = \overline{DA}_{\succ}(P)$ holds.*

Next, we present a condition on the priority structure for the equivalence result. To begin with, *Top-two players*, if any, are the two players who have higher priority than all the others at every object, *i.e.*, i and j in N such that for all $k \in N \setminus \{i, j\}$ and all $a \in O$, both $i \succ_a k$ and $j \succ_a k$ hold. We then have the following.

Definition 4.3. *A priority structure $\succ$ is* reversed *if either one of the following holds:*

(1) *there exist no top-two players;*
(2) *there exist top-two players, $i, j \in N$, and other players, $h, \ell \in N \setminus \{i, j\}$ such that $h \succ_a \ell$ and $\ell \succ_b h$ hold for some $a, b \in O$.*

If $\succ$ is not reversed, it is said to be unreversed*.*

Note that the priority structure is reversed if and only if there exist players $i, j, k \in N$ and objects $a, b \in O$ such that $i \succ_a j \succ_a k$ and $k \succ_b i \vee k \succ_b j$ hold. If the priority structure is unreversed, all the players except for the top-two players have the same priority order across the objects.

Theorem 4.1. *Given a priority structure $\succ$, $\overline{DA}_{\succ}(\cdot)$ is implemented in Nash equilibrium with renegotiation if and only if $\succ$ is unreversed.*

Proof. See Appendix B. □

Theorem 4.1 states that the unreversed priority structure is an equivalence condition under which $\overline{DA}_{\succ}$ is implemented in Nash equilibrium with renegotiation. Its proof roughly goes as follows.

In the proof of the *only-if* part, we take an arbitrary reversed priority structure and construct a preference profile P for which there is a Nash equilibrium with renegotiation P' yielding a final allocation $r(DA_{\succ}(P'); P)$ different from $\overline{DA}_{\succ}(P)$. The construction is similar to, but more comprehensive than, the second example provided in the introduction.

The proof of the *if* part consists of three steps: (i) the true preference profile P constitutes a Nash equilibrium with renegotiation; (ii) $r(DA_{\succ}(P); P) = \overline{DA}_{\succ}(P)$ holds; (iii) the allocation achieved in Nash equilibrium with renegotiation is unique. First, to prove (i), we suppose the contrary, *i.e.*, that there exists a player, denoted i, who can profitably deviate from the true preference P_i, while the others are taking P_{-i}. From the strategy-proofness of the true preference profile in DA without renegotiation, if player i obtains an object $\hat{\omega}_i$ by deviation, player i weakly prefers $DA_{\succ,i}(P)$ to $\hat{\omega}_i$. Therefore, player i needs another player j who exchange their objects. It will be proven that such a cooperator j does *not* exists if the priority structure is unreversed. This leads to a contradiction.

The second step (ii) $r(DA_{\succ}(P); P) = \overline{DA}_{\succ}(P)$ is easily shown.

The third step, the uniqueness of the allocation in Nash equilibrium with renegotiation, is most involved among the three steps. To begin with, we will prove that the top-two players obtain the same objects as in $r(DA_{\succ}(P); P)$ in any Nash equilibrium with renegotiation P'. Next, we will prove that the other players also obtain the same objects as in $r(DA_{\succ}(P); P)$. Suppose the contrary, *i.e.*, $r(DA_{\succ}(P'); P) \neq r(DA_{\succ}(P); P)$. Let i be a player who has the highest priority among the players whose objects are different in these two allocations. Because player i obtains what he/she likes most in $r(DA_{\succ}(P); P)$ except those obtained by the players with higher priority than i, i prefers

$r_i(DA_{\succ}(P); P)$ to $r_i(DA_{\succ}(P'); P)$. Since $r_i(DA_{\succ}(P); P)$ is finally obtained by someone, say, j, who has lower priority than i, player i has an incentive to pretend to be j to obtain $r_i(DA_{\succ}(P); P)$. The actual proof is more involved than this because we have to show that such disguise can be successfully executed.

5. Remarks

Two remarks are in order. The first remark is on an alternative definition of a renegotiation mapping. The second is on priority cycles.

5.1. Alternative definition of renegotiation.

If we use a different renegotiation mapping, Theorem 4.1 would not hold. For example, Maskin and Moore (1999) requires that the final allocation should be Pareto optimal and individually rational, but *not* necessarily a strong core[7]. Then, even if the priority structure $\succ$ is unreversed, $\overline{DA}_{\succ}$ may not be implemented in Nash equilibrium with some renegotiation mapping.

Suppose there are three players, $1, 2, 3$, and three objects, x, y, z. Consider a priority structure that satisfies:

$$1 \succ_x 2 \succ_x 3, \text{ for all } a \in O.$$

Suppose further that the preference profile is given by

$$xP_1yP_1z,$$
$$zP_2yP_2x,$$
$$zP_3yP_3x.$$

Then, for the true preference profile, the DA allocation is (x, z, y), and there will be no further renegotiation. Now, suppose further that the renegotiation mapping $\hat{r}$ satisfies:

$$\hat{r}(z, y, x) = \hat{r}(z, x, y) = (x, y, z).$$

Note that this mapping satisfies Pareto optimality and individual rationality, but (x, y, z) is *not* a strong core for (z, x, y).

In this situation, the following profile is a Nash equilibrium with renegotiation:

- player 1 misreports by putting z on the top of the list;
- player 2 and 3 submit the truthful report.

Then, the outcome of DA before renegotiation is (z, y, x), and the final allocation after renegotiation becomes (x, y, z). Because players 1 and 3 obtain what they like most, they have no incentive to deviate. Player 2 has no incentive to deviate, either, because what player 2 can do is to obtain z instead of y before renegotiation, but it would not change the final allocation.

In our setting, $r(z, x, y)$ has to be a strong core for (z, x, y), and therefore, $r(z, x, y) = (x, z, y)$ must hold instead of (x, y, z). If this is the renegotiation mapping, player 2 has an incentive to deviate to obtain x in the beginning so that (x, y, z) cannot be the final allocation of a Nash equilibrium.

---

[7]See also Maskin and Sjöström (2002).

5.2. **Priority cycles.** The unreversedness of priority structure is a sufficient condition for acyclicity as defined by Ergin (2002). The priority structure $\succ$ is said to be *cyclical* if there exist three players i, j, k and two objects a, b such that $i \succ_a j \succ_a k \succ_b i$ holds. The priority structure is acyclical if it is not cyclical.

Ergin (2002) shows that if the priority structure is acyclical, then the DA allocation is Pareto optimal[8], which implies $r(DA_{\succ}(P); P) = DA_{\succ}(P)$. This is because there is no room for renegotiation if the allocation is Pareto optimal. In the present context, acyclicity implies that for all P, for any player i, and any report P'_i, we have

$$r_i(DA_{\succ}(P); P)\ R_i\ DA_{\succ,i}(P'_i, P_{-i}).$$

Of course, this does not necessarily imply that $\overline{DA}_{\succ}$ is implemented in Nash equilibrium with renegotiation. In fact, the second example in the introduction is the case of acyclical but reverse priority structure. Unreversedness, a stronger concept than acyclicity, plays an important role for Nash implementation with renegotiation.

6. Conclusion

The present paper studies DA with renegotiation. We have two main results. The first theorem states that DA is not strategy-proof with renegotiation for *any* priority structure. Then, the second theorem states that $\overline{DA}_{\succ}$ is implemented in Nash equilibrium with renegotiation if and only if the priority structure is unreversed.

The present paper examines DA with renegotiation, which, as a matter of course, is restrictive from the viewpoint of the literature on mechanism design with renegotiation. The restriction is twofold. First, the social choice rule considered in the model is the DA rule denoted by $\overline{DA}_{\succ}$ rather than a general rule. Second, the mechanism is restricted to $DA_{\succ}$. If we generalize the rule and the mechanism, there is no guarantee that we obtain a tight result, *i.e.*, a necessary and sufficient condition for implementability in Nash equilibrium.

DA has been successfully applied to school choice systems and medical residency matching programs (see Roth, 2008). The present paper suggests that this success may be because renegotiation is either prohibited or costly in these applications. In the application of DA to other cases, we have to examine them carefully to see if allocations are subject to renegotiation.

[8]Ergin (2002) shows the equivalence result.

Appendix A. Proof of Theorem 3.1

Proof. Take any N with $|N| \geq 3$ and O with $|O| \geq |N|$. Write $N = \{1, 2, 3, \dots, n\}$ and $O = \{a_1, a_2, a_3, \dots, a_m\}$ with $m \geq n$.

We assume without loss of generality that $1 \succ_{a_1} 2 \succ_{a_1} 3$ holds. Let P satisfy the following: for players $1, 2, 3$ and for all $k > 3$:

$$a_3 P_1 a_2 P_1 a_1 P_1 a_k$$
$$a_1 P_2 a_2 P_2 a_3 P_2 a_k$$
$$a_1 P_3 a_2 P_3 a_3 P_3 a_k;$$

for all $j > 3$, for all $k \neq j$, $a_j P_j a_k$ holds. We divide into two cases in terms of $\succ_{a_2}$.

Case (I): $2 \succ_{a_2} 3$. In this case, $r(DA_{\succ}(P); P) = (a_3, a_1, a_2, a_4, \dots, a_n)$. Suppose that P'_1 puts a_1 on the top of the preference order for player 1. Then, we have

$$DA_{\succ}(P'_1, P_{-1}) = (a_1, a_2, a_3, a_4, \dots, a_n).$$

After renegotiation, we have

$$r(DA_{\succ}(P'_1, P_{-1}); P) = (a_3, a_2, a_1, a_4, \dots, a_n).$$

Let P''_2 be a preference order that puts a_3 on the top. Then, we have

$$DA_{\succ}(P''_2, P'_1, P_{-\{1,2\}}) = (a_1, a_3, a_2, a_4, \dots, a_n).$$

After renegotiation, we have

$$r(DA_{\succ}(P''_2, P'_1, P_{-\{1,2\}}); P) = (a_3, a_1, a_2, a_4, \dots, a_n).$$

Thus, we have

$$r_2(DA_{\succ}(P''_2, P'_1, P_{-\{1,2\}}); P) = a_1 \ P_2 \ r_2(DA_{\succ}(P'_1, P_{-1}); P) = a_2.$$

Hence, DA with renegotiation is not strategy-proof in this case.

Case (II): $3 \succ_{a_2} 2$. In this case, $r(DA_{\succ}(P); P) = (a_3, a_1, a_2, a_4, \dots, a_n)$.

Suppose that P'_1 puts a_1 on the top of the preference order for player 1. Then, we have

$$DA_{\succ}(P'_1, P_{-1}) = (a_1, a_3, a_2, a_4, \dots, a_n).$$

After renegotiation, we have

$$r(DA_{\succ}(P'_1, P_{-1}); P) = (a_3, a_1, a_2, a_4, \dots, a_n).$$

Let P''_3 be a preference order that puts a_3 on the top. Then, we have

$$DA_{\succ}(P''_3, P'_1, P_{-\{1,3\}}) = (a_1, a_2, a_3, a_4, \dots, a_n).$$

After renegotiation, we have

$$r(DA_{\succ}(P''_3, P'_1, P_{-\{1,3\}}); P) = (a_3, a_2, a_1, a_4, \dots, a_n).$$

Thus, we have

$$r_3(DA_{\succ}(P''_3, P'_1, P_{-\{1,3\}}); P) = a_1 \ P_3 \ r_3(DA_{\succ}(P'_1, P_{-1}); P) = a_2.$$

Hence, DA with renegotiation is not strategy-proof in this case, either.

□

Appendix B. Proof of Theorem 4.1

B.1. **Unreversed Priority and Top Trading Cycle (TTC).** Before proving Theorem 4.1, we need some claims, which will be presented in this subsection of the appendix. If we consider an allocation of DA as an endowment, the renegotiation process is closely related to the market without money. Therefore, it is convenient to introduce the auxiliary market to make some statements. Let $O^{\omega} = \{a \in O | \; \exists i \in N \omega_i = a\}$.

Definition B.1. *Given $P \in \mathcal{P}$ and $\omega \in A$, $(p, \mu) \in \mathbb{R}_+^{O^{\omega}} \times A$ is a* market equilibrium *w.r.t. (P, ω) if*

- *$\mu_i R_i a$ holds for all $i \in N$ and all $a \in O^{\omega}$ with $p_a \leq p_{\omega_i}$,*
- *μ is feasible w.r.t. ω,*

where $p = (p_a)_{a \in O^{\omega}} \in \mathbb{R}_+^{O^{\omega}}$ is a price vector.

Shapley and Scarf (1974) shows that for any initial endowment profile, a market equilibrium (p, μ) exists.

Lemma B.1. *(Shapley & Scarf, 1974) For all $P \in \mathcal{P}$ and $\omega \in A$, a market equilibrium exists w.r.t. (P, ω).*

Given the initial endowment profile $\omega \in A$ of the second stage, transactions occur in the form of trading cycles (Gale's top trading cycle). A *trading cycle* is given by

$$((i_1, \omega_{i_1}), (i_2, \omega_{i_2}), \ldots, (i_K, \omega_{i_K})), \text{ with } (i_1, \omega_{i_1}) = (i_K, \omega_{i_K}),$$

where i_{k-1} sells $\omega_{i_{k-1}}$ to i_{k-2} and buys ω_{i_k} from i_k ($k = 2, \ldots, K$). A trading cycle may contain only one player. In such a cycle, the player does not actually trade his/her object with another player.

We have the following unique existence result due to Roth and Postlewaite (1977).

Lemma B.2. *(Roth & Postlewaite, 1977) For all $P \in \mathcal{P}$ and $\omega \in A$, a market equilibrium object allocation of $\mu(\omega)$ uniquely exists w.r.t. (P, ω).*

The following corollary will be used in the sequel.

Corollary B.3. *For any $P, P' \in \mathcal{P}$, if $\mu = r(DA_{\succ}(P'); P)$ then there exists a price vector $p \in \mathbb{R}_+^{\omega}$ such that (p, μ) is a market equilibrium w.r.t. $(P, DA_{\succ}(P'))$.*

Take $P \in \mathcal{P}$ and $\omega \in A$ as given. Suppose that (p, μ) is a market equilibrium w.r.t. (P, ω). Partition O^{ω} into $O_1, O_2, \ldots, O_T$ in such a way that for all $t = 1, \ldots, T$ and all $a.b \in O_t$, $p_a = p_b$ holds, and that for all $t < t'$, $a \in O_t$, and $b \in O_{t'}$, $p_a > p_b$ holds. In other words, O_t's form an equivalence class of objects in terms of equilibrium price. Let $S_t = \{i \in N | \; \omega_i \in O_t\}$. Given $i \in N$, let $S(i) \subset N$ be the set of players who are in the same trading cycle as player i. It is verified that $S(i)$'s form an equivalence class. In particular, if $j \in S(i)$ holds, then we have $S(i) = S(j)$. It is also verified that $S(i)$'s form a finer partition of N than S_t's, *i.e.*, $S(i) \cap S_t \neq \emptyset$ implies $S(i) \subset S_t$.

We are now in a position to state the following lemma, which will be used in the proof of Theorem 4.1. In the statement, we write $k \succ i$ if $k \succ_a i$ holds for all $a \in O$.

Lemma B.4. *Assume that the priority structure is unreversed. Take $P \in \mathcal{P}$ as given. Let P' be a Nash equilibrium with renegotiation w.r.t. P, and $\omega' = DA_{\succ}(P')$. Also, let (p, μ') be a market equilibrium w.r.t. (P, ω'). Then, for all $t = 1, \ldots T$, for all $i \in S_t$, and for all $j \in S(i)$, there exists no $k \succ i$ such that $\omega'_j P_k \mu'_k$ holds.*

Proof. We prove this lemma by induction. Suppose $i \in S_1$. In TTC, it is obvious that $\mu'_i P_i \omega'_\ell$ holds for any $\ell \in S_\tau$ for any $\tau = 1, \ldots, T$. Suppose the contrary, *i.e.*, there exist $i \in S_1$, $j \in S(i)$, and $k \in S_\tau$ ($\tau = 1, \ldots, T$) such that $k \succ i$ and $\omega'_j P_k \mu'_k$ hold. If $\tau = 1$, then k must have designated ω'_j or better, which is strictly better than μ'_k. Therefore, we have $\tau > 1$.

Then we can construct a trading cycle:

$$(j_0, \omega_{j_0}), (j_1, \omega_{j_1}), \ldots, (j_m, \omega_{j_m}), \ldots$$

such that $j_0 = j$, $j_{m'} \succ k$ for all $m' = 0, 1, \ldots, m-1$, and $k \succ j_m$[9]. Then k can obtain ω'_{j_m} without affecting $\omega'_{j_{m'}}$ with $m' < m$. Then, player k can obtain ω'_j in this trading cycle.

Suppose we have proven the claim up to $t - 1 = 1, \ldots, T-1$. We would like to show the claim for t. Suppose the contrary, *i.e.*, there exist $i \in S_t$, $j \in S(i)$, and $k \in S_\tau$ ($\tau = 1, \ldots, T$) such that $k \succ i$ and $\omega'_j P_k \mu'_k$ hold. If $\tau \leq t$, then k must have designated ω'_j or better, which is strictly better than μ'_k. Therefore, we have $\tau > t$.

Then we can construct a trading cycle:

$$(j_0, \omega_{j_0}), (j_1, \omega_{j_1}), \ldots, (j_m, \omega_{j_m}), \ldots$$

such that $j_0 = j$, $j_{m'} \succ k$ for all $m' = 0, 1, \ldots, m-1$, and $k \succ j_m$. Then k can obtain ω'_{j_m} without affecting $\omega'_{j_{m'}}$ with $m' < m$. Indeed, since we have proven the claim up to $t-1$, the only cycles that may be affected by k's deviation are the ones of which objects are not preferred to $\mu'_{j_{m'}}$ by $j_{m'}$ for $m' = 0, \ldots, m-1$. Then, player k can obtain ω'_j in this trading cycle. □

B.2. **Proof of Theorem 4.1.**

Proof.
[*only-if* part]

Assume that the priority structure is reversed. Take any N with $|N| \geq 3$ and O with $|O| \geq |N|$. Write $N = \{1, 2, 3, \ldots, n\}$ and $O = \{a_1, a_2, a_3, \ldots, a_m\}$ with $m \geq n$. We assume without loss of generality that $1 \succ_{a_1} 2 \succ_{a_1} 3$ holds. Since $\succ$ is reversed, we may further assume without loss of generality that either $3 \succ_{a_3} 2$ or $3 \succ_{a_3} 1$ (or both) holds.
Case (I): $3 \succ_{a_3} 2$.

Let P satisfy the following: for players $1, 2, 3$ and for all $k > 3$:

$$a_3 P_1 a_2 P_1 a_1 P_1 a_k$$
$$a_1 P_2 a_2 P_2 a_3 P_2 a_k$$
$$a_1 P_3 a_2 P_3 a_3 P_3 a_k;$$

for all $j > 3$, for all $k \neq j$, $a_j P_j a_k$ holds. Then, we have $r(DA_\succ(P); P) = (a_3, a_1, a_2, a_4, \ldots, a_n)$. Suppose that P'_1 puts a_1 on the top of the preference order for player 1, and that P'_3 puts a_3 on the top of the preference order for player 3. Then, we have

$$DA_\succ(P'_1, P'_3, P_{-1,3}) = (a_1, a_2, a_3, a_4, \ldots, a_n).$$

After renegotiation, we have

$$r(DA_\succ(P'_1, P'_3, P_{-1,3}); P) = (a_3, a_2, a_1, a_4, \ldots, a_n).$$

Since nobody has an incentive to deviate, $(P'_1, P'_3, P_{-1,3})$ is a Nash equilibrium, but

$$r(DA_\succ(P'_1, P'_3, P_{-1,3}); P) \neq r(DA_\succ(P); P) = (a_3, a_1, a_2, a_4, \ldots, a_n).$$

[9]We write $i \succ j$ if $i \succ_a j$ for all $a \in O$.

Case (II): $3 \succ_{a_3} 1$.

There are three subcases, $3 \succ_{a_3} 1 \succ_{a_3} 2$, $3 \succ_{a_3} 2 \succ_{a_3} 1$, and $2 \succ_{a_3} 3 \succ_{a_3} 1$. But, the first two subcases have already been taken care of in Case (I). Therefore, we consider the case of $2 \succ_{a_3} 3 \succ_{a_3} 1$.

Let P satisfy the following: for players $1, 2, 3$ and for all $k > 3$:

$$a_3 P_1 a_2 P_1 a_1 P_1 a_k$$
$$a_1 P_2 a_2 P_2 a_3 P_2 a_k$$
$$a_3 P_3 a_2 P_3 a_1 P_3 a_k;$$

for all $j > 3$, for all $k \neq j$, $a_j P_j a_k$ holds. Then, we have $r(DA_{\succ}(P); P) = (a_2, a_1, a_3, a_4, \ldots, a_n)$. Suppose that P'_1 puts a_1 on the top of the preference order for player 1, and that P'_2 puts a_3 on the top of the preference order for player 2. Then, we have

$$DA_{\succ}(P'_1, P'_2, P_{-1,2}) = (a_1, a_3, a_2, a_4, \ldots, a_n).$$

After renegotiation, we have

$$r(DA_{\succ}(P'_1, P'_2, P_{-1,2}); P) = (a_3, a_1, a_2, a_4, \ldots, a_n).$$

Since nobody has an incentive to deviate, $(P'_1, P'_2, P_{-1,2})$ is a Nash equilibrium, but

$$r(DA_{\succ}(P'_1, P'_2, P_{-1,2}); P) \neq r(DA_{\succ}(P); P) = (a_2, a_1, a_3, a_4, \ldots, a_n).$$

[*if* part]

Assume that the priority structure is unreversed. Assume without loss of generality that $\{1, 2\}$ be the set of the top-two players, and that $N \setminus \{1, 2\} = \{3, 4, \ldots, n\}$ is aligned in such a way that $k \succ_a k + 1$ holds for all $a \in O$ and all $k = 3, \ldots, n - 1$.

First, it is shown that the true profile P is a NE. Suppose the contrary, *i.e.*, the true profile is not a NE. Then, there is a player i who is better off by deviation when all the other players submit the true profile. Let P denote the profile of the true preferences and P'_i denote the deviation of player i. Let $\omega = DA_{\succ}(P) = \overline{DA}_{\succ}(P)$, $\mu = r(DA_{\succ}(P); P)$, $\hat{\omega} = DA_{\succ}(P'_i, P_{-i})$ and $\hat{\mu} = r(DA_{\succ}(P'_i, P_{-i}); P)$. Due to the strategy-proofness of DA, i obtains $\hat{\omega}_i$ such that $\omega_i R_i \hat{\omega}_i$. Since i is better off, $\hat{\mu}_i P_i \mu_i R_i \hat{\omega}_i$ also holds. We divide into two cases.

Case1: Suppose that i is one of the top-two players, say, 1. Since $\hat{\mu}_1 \neq \mu_1$ holds, $\hat{\mu}_1 = \mu_2$ and $2 \succ_{\hat{\mu}_1} 1$ must hold: otherwise, player 1 could have obtained $\hat{\mu}_1$ for the true preference. Moreover, $\hat{\mu}_1$ must be the best object for player 2. Therefore, player 2 would never give $\hat{\mu}_1$ to player 1. A contradiction.

Case2: Suppose that i is not one of the top-two players. Because of the unreversedness, for all $k < i$ and $j \geq i$, $\hat{\omega}_k = \omega_k$ and $\omega_k P_k \hat{\omega}_j$ hold, and for all $j > i$, $\omega_i P_i \omega_j$ holds. Therefore, no $k < i$ is interested in $\hat{\omega}_j$ with $j \geq i$. Thus, no $k < i$ receives an object from any $j \geq i$ through renegotiation. Hence, i cannot receive an object from any $k < i$. Player i cannot be better off.

Second, we would like to show $r(DA_{\succ}(P); P) = \overline{DA}_{\succ}(P)$, or $\mu = \omega$, *i.e.*, there is no renegotiation for the true profile. Suppose not. Then there exists S that blocks ω. Let i be the smallest number in S. If i is either 1 or 2, then using the same argument as above, we have a contradiction. Suppose $i \geq 3$. Again, for all $j > i$, $\omega_i P_i \omega_j$ holds. Therefore, $\mu_i = \omega_i$ holds. Thus, coalition $S \setminus \{i\}$ should be able to block ω. Repeating this argument, we conclude that an empty set should be able to block ω, which is a contradiction.

Third, consider a profile P' that is a Nash equilibrium with renegotiation. Let $\omega' = DA_{\succ}(P')$ and $\mu' = r(DA_{\succ}(P'); P)$. We would like to show $r(DA_{\succ}(P'); P) = \overline{DA_{\succ}}(P)$, or equivalently $\mu' = \omega (= \mu)$.

We show this statement in two steps. To begin with, we would like to show $\mu'_i = \mu_i$ for $i = 1, 2$. Suppose not, *i.e.*, say, $\mu'_1 \neq \mu_1$. If $\mu'_1 P_1 \mu_1$ holds, then player 2 must have taken μ'_1 for P, *i.e.*, $\mu_2 = \mu'_1$ and $2 \succ_{\mu_2} 1$. This implies that player 2 prefers μ_2 to any other object. Therefore, player 2 never gives it to player 1. A contradiction. Thus, $\mu_1 P_1 \mu'_1$ holds. Since μ'_1 is at least second best for player 1, μ_1 is the most preferred object for player 1, $2 \succ_{\mu_1} 1$, and $\omega'_2 = \mu_1$ hold. Since player 2 did not take μ_1 for P, μ_2 must be the most preferred object for player 2. Then by obtaining μ_2 in DA, player 1 should be able to exchange it with 2 for μ_1 with renegotiation. A contradiction.

Next, we would like to show $\mu'_i = \mu_i$ for all $i \geq 3$. Let i be the smallest number such that $\mu'_i \neq \mu_i$ holds, *i.e.*, $\mu'_k = \mu_k$ for all $k < i$. This implies $\mu_i P_i \mu'_i$. If $\omega'_k \neq \omega_i$ holds for all $k < i$, then player i can obtain at least ω_i, which implies $\mu'_i P_i \omega_i = \mu_i$. This is impossible. Therefore, there exists $k_0 < i$ with $\omega'_{k_0} = \omega_i \neq \omega_{k_0}$.

Corollary B.3 implies that (p, μ') is a market equilibrium w.r.t. (P, ω') for some p. We then have a partition of N, $S_1, \ldots, S_T$, such that for any $t < t'$, any $a \in S_t$, and any $b \in S_{t'}$, $p_a > p_b$ holds. Then there exist S_τ that contains k_0 and $S_{\tau'}$ with $\tau' > \tau$ that contains i. There must be a sequence

$$(k_0, \omega'_{k_0}), (k_1, \omega'_{k_1}), \ldots, (k_m, \omega'_{k_m})$$

such that $\omega'_{k_\ell} = \mu'_{k_{\ell-1}}$ for $\ell = 1, \ldots, m$, and $\mu'_{k_m} = \omega_i$. Note that $m > i$ holds since $\mu'_k = \omega_k$ holds for all $k < i$, and $\mu'_i \neq \omega_i$. But, we have $\omega_i P_i \mu'_i$. This is a contradiction to Lemma B.4.

□

References

Dubins, L. E., & Freedman, D. A. (1981). Machiavelli and the Gale-Shapley algorithm. *American Mathematical Monthly*, *88*(7), 485–494.

Ergin, H. I. (2002). Efficient resource allocation on the basis of priorities. *Econometrica*, *70*(6), 2489–2497.

Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, *69*(1), 9–15.

Hatfield, J. W., & Kojima, F. (2010). Substitutes and stability for matching with contracts. *Journal of Economic Theory*, *145*(5), 1704–1723.

Kaneko, M. (1982). The central assignment game and the assignment markets. *Journal of Mathematical Economics*, *10*(2-3), 205–232.

Maskin, E., & Moore, J. (1999). Implementation and renegotiation. *Review of Economic Studies*, *66*(1), 39–56.

Maskin, E., & Sjöström, T. (2002). Implementation theory. *Handbook of social Choice and Welfare*, *1*, 237–288.

Roth, A. E. (1982). The economics of matching: Stability and incentives. *Mathematics of Operations Research*, *7*(4), 617–628.

Roth, A. E. (2008). Deferred acceptance algorithms: history, theory, practice, and open questions. *International Journal of Game Theory*, *36*(3), 537–569.

Roth, A. E., & Postlewaite, A. (1977). Weak versus strong domination in a market with indivisible goods. *Journal of Mathematical Economics*, *4*(2), 131–137.

Shapley, L., & Scarf, H. (1974). On cores and indivisibility. *Journal of Mathematical Economics*, *1*(1), 23–37.

Sotomayor, M. (2008). The stability of the equilibrium outcomes in the admission games induced by stable matching rules. *International Journal of Game Theory*, *36*(3/4), 621–640.

Wako, J. (1984). A note on the strong core of a market with indivisible goods. *Journal of Mathematical Economics*, *13*(2), 189–194.

Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.
E-mail address: amatsui@e.u-tokyo.ac.jp
URL: http://www.e.u-tokyo.ac.jp/~amatsui

Department of Economics, Northwestern University, 2211 Campus Drive, Evanston, IL 60208, U.S.A.