IMPOSSIBILITY OF STABLE AND NONBOSSY MATCHING MECHANISM

FUHITO KOJIMA

Cowles Foundation, Yale University
30 Hillhouse Avenue, New Haven, CT 06510
fuhitokojima1979@gmail.com

Abstract. Stability is a central concept in matching theory, while nonbossiness is important in many allocation problems. We show that these properties are incompatible: There does not exist a matching mechanism that is both stable and nonbossy. *JEL Classification Numbers*: C70, D61, D63.

Keywords: matching, stability, nonbossiness, impossibility theorem.

1. Introduction

Initiated by Gale and Shapley (1962), matching theory has influenced the design of labor markets and student assignment systems.[1] Stability plays a central role in the theory: A matching is stable if there is no individual agent who prefers being unmatched to being assigned to her allocation in the matching, and there is no pair of agents who prefer being assigned to each other to being assigned to their respective allocations in the matching. In real-world applications, empirical studies have shown that stable mechanisms often succeed whereas unstable ones often fail.[2]

The concept of nonbossiness (Satterthwaite and Sonnenschein 1981) is important in many allocation problems. A mechanism is nonbossy if an agent cannot change allocation of other agents without changing her own allocation. Normatively, the concept requires a form of fairness: It is arguably unfair for an agent to be affected by changes of reported preferences of someone else, even though the change has no consequence on the allocation of the latter. Also, if an allocation violates nonbossiness, then it may invite strategic

Date: January 21, 2009.

[1]For a survey of this theory, see Roth and Sotomayor (1990). For applications to labor markets, see Roth (1984) and Roth and Peranson (1999). For applications to student assignment, see for example Abdulkadiroğlu and Sönmez (2003), Abdulkadiroğlu, Pathak, Roth, and Sönmez (2005) and Abdulkadiroğlu, Pathak, and Roth (2005).

[2]For a summary of this evidence, see Roth (2002).

manipulation: an agent affected by another might pay a small transfer to the latter in return to reporting preferences that results in a preferable allocation to him. As the latter agent may not be affected by changing her own reported preferences, she may well agree to engage in such manipulations.

Given the importance of nonbossiness, the concept has been studied extensively in the context of indivisible good allocations. In that environment, the combination of strategy-proofness and nonbossiness is equivalent to the group strategy-proofness, and allocation mechanisms that are efficient and group strategy-proof have been studied and characterized by Papai (2000) and Pycia and Unver (2009). Ergin (2002) characterizes the market structures in which the student-proposing deferred acceptance algorithm (Gale and Shapley 1962) is nonbossy and, since that mechanism is strategy-proof, group strategy-proof.[3]

Although these two properties are important properties, we show that these properties are incompatible: There does not exist a matching mechanism that is both stable and nonbossy. Thus any stable mechanism can cause an undesirable consequence where an agent influences allocation of other agents without changing her own allocation.

2. Model

A (one-to-one) matching problem is tuple $(S, C, \succ)$. S and C are finite and disjoint sets of students and colleges. For each student $s \in S$, $\succ_s$ is a strict preference relation over C and being unmatched (being unmatched is denoted by $\emptyset$). For each college $c \in C$, $\succ_c$ is a strict preference relation over S and being unmatched, $\emptyset$. We write $\succ = (\succ_i)_{i \in S \cup C}$. A **matching** is a vector $\mu = (\mu_s)_{s \in S}$ assigning a college $\mu_s \in C$ or $\emptyset$ to each student s, where at most one student is assigned to college c. We write $\mu_c = s$ if and only if $\mu_s = c$ and $\mu_c = \emptyset$ if there is no s with $\mu_s = c$.

We say that matching μ is **blocked** by $(s, c) \in S \times C$ if $c \succ_s \mu_s$ and $s \succ_c \mu_c$. A matching μ is **individually rational** if $\mu_i \succ_i \emptyset$ for every $i \in S \cup C$. A matching μ is **stable** if it is individually rational and is not blocked.

A **mechanism** is a function φ from the set of preference profiles to the set of matchings. Mechanism φ is **stable** if $\varphi(\succ)$ is a stable matching for every preference profile. Existence

[3]When the market structure does not satisfy Ergin's condition, only a weaker version of group strategy-proofness holds (and the mechanism violates nonbossiness). That is, no group of students can make each of its members strictly better off by jointly misreporting their preferences. This latter result is first shown by Dubins and Freedman (2002) and extended by Martinez, Masso, Neme, and Oviedo (2004), Hatfield and Kojima (2007) and Hatfield and Kojima (2008).

of a stable mechanism is shown by Gale and Shapley (1962). They propose deferred acceptance algorithms, which find stable matchings for any preference profile.

3. Result

We introduce the concept of nonbossiness (Satterthwaite and Sonnenschein 1981).

Definition 1. A mechanism φ is **nonbossy** if, for any $\succ$ and $\succ'_i$, $\varphi_i(\succ'_i, \succ_{-i}) = \varphi_i(\succ)$ implies $\varphi(\succ'_i, \succ_{-i}) = \varphi(\succ)$.

In words, a mechanism is nonbossy if an agent cannot change allocation of other agents unless doing so also changes her own allocation. With this concept, we now proceed to present the following impossibility result.

Theorem 1. There does not exist a mechanism that is stable and nonbossy.

Proof. Consider a problem where $C = \{c_1, c_2, c_3\}, S = \{s_1, s_2, s_3\}$, and preferences are given by

$$\succ_{c_1} : s_1, s_2, s_3, \emptyset,$$

$$\succ_{c_2} : \emptyset,$$

$$\succ_{c_3} : s_3, s_2, s_1, \emptyset,$$

$$\succ_{s_1} : c_3, c_2, c_1, \emptyset,$$

$$\succ_{s_2} : c_3, c_2, c_1, \emptyset$$

$$\succ_{s_3} : c_1, c_2, c_3, \emptyset,$$

where $\succ_{c_1}: s_1, s_2, s_3, \emptyset$, means "according to preferences $\succ_{c_1}$ of c_1, s_1 is most preferred and followed by s_2, s_3 and $\emptyset$ in this order," for example. There exists a unique stable matching $\varphi(\succ)$ given by

$$\varphi(\succ) = \begin{pmatrix} c_1 & c_2 & c_3 & \emptyset \\ s_1 & \emptyset & s_3 & s_2 \end{pmatrix},$$

which means that c_1 is matched to s_1, c_3 is matched to s_3, and c_2 and s_2 are unmatched. Consider $\succ'_{s_2}$ given by

$$\succ'_{s_2}: \emptyset.$$

Now there are two stable matchings, μ and μ', given by

$$\mu = \begin{pmatrix} c_1 & c_2 & c_3 & \emptyset \\ s_3 & \emptyset & s_1 & s_2 \end{pmatrix},$$

and

$$\mu' = \begin{pmatrix} c_1 & c_2 & c_3 & \emptyset \\ s_1 & \emptyset & s_3 & s_2 \end{pmatrix},$$

respectively. Now consider the following two cases.

First, consider the Case in which $\varphi(\succ'_{s_2}, \succ_{-s_2}) = \mu$. In that case apparently we have $\varphi_{s_2}(\succ'_{s_2}, \succ_{-s_2}) = \varphi_{s_2}(\succ)$ and $\varphi(\succ'_{s_2}, \succ_{-s_2}) \neq \varphi(\succ)$, thus φ is not nonbossy.

Second, consider the Case in which $\varphi(\succ'_{s_2}, \succ_{-s_2}) = \mu'$. Now consider $\succ''_{c_2}$ given by

$$\succ''_{c_2} : s_1, s_2, s_3.$$

Then $\varphi(\succ''_{c_2}, \succ'_{s_2}, \succ_{-c_2,s_2})$ is given by

$$\varphi(\succ''_{c_2}, \succ'_{s_2}, \succ_{-c_2,s_2}) = \begin{pmatrix} c_1 & c_2 & c_3 & \emptyset \\ s_3 & \emptyset & s_1 & s_2 \end{pmatrix}.$$

Therefore we have that $\varphi_{c_2}(\succ''_{c_2}, \succ'_{s_2}, \succ_{-c_2,s_2}) = \varphi_{c_2}(\succ'_{s_2}, \succ_{-s_2})$ and $\varphi(\succ''_{c_2}, \succ'_{s_2}, \succ_{-c_2,s_2}) \neq \varphi(\succ'_{s_2}, \succ_{-s_2})$, so φ is not nonbossy. This completes the proof. □

As mentioned in the Introduction, stability and nonbossiness are regarded as important properties of allocation mechanisms. However, Theorem 1 shows that these desiderata are incompatible. Thus stable mechanisms cannot avoid the situation where an agent influences allocation of other agents without changing her own allocation.

References

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005): "The New York City High School Match," *American Economic Review Papers and Proceedings*, 95, 364–367.

Abdulkadiroğlu, A., P. A. Pathak, A. E. Roth, and T. Sönmez (2005): "The Boston Public School Match," *American Economic Review Papers and Proceedings*, 95, 368–372.

Abdulkadiroğlu, A., and T. Sönmez (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

Dubins, L. E., and D. A. Freedman (2002): "Machiavelli and the Gale-Shapley algorithm," *American Mathematical Monthly*, 70, 2489–2497.

Ergin, H. (2002): "Efficient Resource Allocation on the Basis of Priorities," *Econometrica*, 88, 485–494.

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

Hatfield, J. W., and F. Kojima (2007): "Group Incentive Compatibility for Matching with Contracts," forthcoming, *Games and Economic Behavior*.

——— (2008): "Substitutes and Stability for Matching with Contracts," mimeo.

Martinez, R., J. Masso, A. Neme, and J. Oviedo (2004): "On group strategy-proof mechanisms for a many-to-one matching model," *International Journal of Game Theory*, 33, 115–128.

Papai, S. (2000): "Strategyproof Assignment by Hierarchical Exchange," *Econometrica*, 68, 1403–1433.

Pycia, M., and U. Unver (2009): "A Theory of House Allocation and Exchange Mechanisms," mimeo.

Roth, A. E. (1984): "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of Political Economy*, 92, 991–1016.

——— (2002): "The economist as engineer: Game theory, experimentation, and computation as tools for design economics," *Econometrica*, 70, 1341–1378.

Roth, A. E., and E. Peranson (1999): "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, 89, 748–780.

Roth, A. E., and M. O. Sotomayor (1990): *Two-sided matching: a study in game-theoretic modeling and analysis*. Econometric Society monographs, Cambridge.

Satterthwaite, M., and H. Sonnenschein (1981): "Strategy-Proof Allocation Mechanisms at Differentiable Points," 48, 587–597.