On a characterization of stable matchings

Hiroyuki Adachi*

Institute of Economic Research, Kyoto University, Kyoto 606-8501, Japan

Abstract

The set of stable matchings in the Gale–Shapley marriage problem is characterized as the fixed points of an increasing function. Its well-known non-emptiness and lattice property are an immediate consequence of Tarski's fixed point theorem.   © 2000 Elsevier Science S.A. All rights reserved.

Keywords: Stable matching; Two-sided matching; Gale–Shapley marriage problem; Lattice

JEL classification: C71

1. Introduction

This note presents an alternative way to characterize the set of stable matchings in the Gale–Shapley marriage problem with agents having strict preferences. When agents have strict preferences, we will show that the set of stable matchings coincides with the set of fixed points of a certain mapping. By recognizing that the mapping possesses monotonicity, the lattice property of the set of stable matchings, as well as its non-emptiness, is proved as an immediate implication of Tarski's fixed point theorem. When preferences are not strict, such a formulation cannot fully characterize the set of stable matchings, but can provide further proof that every marriage problem has a stable matching.

This note is organized as follows. In Section 2 we give a brief introduction to the Gale–Shapley marriage problem and present a key observation which leads to our formulation. Section 3 formalizes our idea and shows that our formulation fully characterizes the set of stable matchings when preferences are strict. Proofs to the previously known results mentioned above immediately follow from this formulation. In Section 4, by introducing tie-breaking rules in not-necessarily-strict preferences, our formulation is shown to give further proof to the existence of a stable matching when preferences are not necessarily strict. The discussion follows in Section 5.

*Tel.: +81-75-753-7136; fax: +81-75-753-7198.
E-mail address: adachi@kier.kyoto-u.ac.jp (H. Adachi)

2. Preliminary

This section gives a brief description of Gale–Shapley marriage problems, following the exposition by Roth and Sotomayor (1990).

In a marriage market, there are two finite and disjoint sets M and W: $M = \{m_1, m_2, \ldots, m_n\}$ is the set of men, and $W = \{w_1, w_2, \ldots, w_k\}$ is the set of women. Each man has preferences over the women and each woman has preferences over the men. An agent may prefer to remain single rather than get married to some agent of the opposite sex. So a typical man m's preference ordering $>_m$ is represented by an ordered list over the set $W \cup \{m\}$. We assume preferences are rational. We denote $w >_m w'$ to mean m prefers w to w', and $w \geq_m w'$ to mean m likes w at least as much as w'. We also write $w =_m w'$ to mean man m is indifferent between mating w and w', and write $w = w'$ to mean w is the same person as w'. Woman w is *acceptable* to man m if he likes her at least as much as staying single, i.e. if $w \geq_m m$. An individual is said to have *strict* preferences if he or she is not indifferent between any two acceptable alternatives. We restrict the domain of preferences except in Section 4:

Assumption 2.1. *Every individual has strict preferences.*

This assumption brings great simplification to the problem; we can use the indifference statement $w =_m w'$ and the identity statement $w = w'$ interchangeably (if w and w' are acceptable to man m).

An outcome of the marriage market is described by a rule that matches an agent to an agent:

Definition 2.2. A *matching* μ is a one-to-one correspondence from the set $M \cup W$ onto itself of order two (that is, $\mu^2(x) = x$) such that if $\mu(m) \neq m$ then $\mu(m) \in W$ and if $\mu(w) \neq w$ then $\mu(w) \in M$.

Definition 2.3. A matching μ is *stable* if it satisfies

(IR) $\mu(m) \geq_m m$, $\forall m \in M$ and $\mu(w) \geq_w w$, $\forall w \in W$; i.e. a matching μ is individually rational.

(S) $\nexists(m,w)$ such that $w >_m \mu(m)$ and $m >_w \mu(w)$; i.e. a matching μ is not blocked by any pair of man and woman.

Lemma 2.4. *Let the assumption of strict preferences hold. Suppose a matching μ satisfies condition (IR). Then μ satisfies condition (S) iff it meets the following condition*:

(S′) $\nexists(m,w)$ *such that* $\{w >_m \mu(m)$ *and* $m \geq_w \mu(w)\}$ *or* $\{w \geq_m \mu(m)$ *and* $m >_w \mu(w)\}$.

Proof. Omitted. □

Example 2.5. ((Example 2.4 in Roth and Sotomayor, 1990)) $M = \{m_1, m_2, m_3\}$ and $W = \{w_1, w_2, w_3\}$ have the following preferences:

$$P(m_1): w_2, w_1^*, w_3, m_1, \quad P(w_1): m_1^*, m_3, m_2, w_1,$$
$$P(m_2): w_1, w_3^*, w_2, m_2, \quad P(w_2): m_3^*, m_1, m_2, w_2,$$
$$P(m_3): w_1, w_2^*, w_3, m_3, \quad P(w_3): m_1, m_3, m_2^*, w_3.$$

Everyone has strict preferences and prefers marrying any one of the opposite sex rather than being

single. Man m_1, for instance, prefers w_2 most, then w_1, and so forth. Matching μ' is one of the two stable matchings where μ' is indicated by *. A key observation that leads to our formulation is that, say, the mate of m_1 under μ', $\mu'(m_1) = w_1$, is the greatest element (with respect to m_1's preferences) among the potential partners who like m_1 at least as much as their mates under μ', $\{w \in W | m_1 \geq_w \mu(w)\} \cup \{m_1\}$. This property holds for every agent. The observation suggests that the set of stable matchings may be characterized by the solutions to a set of individual maximization problems. This turns out to be true when preferences are strict.

Gale and Shapley (1962) show that for every marriage problem there exists a stable matching, and if preferences are strict there exist *M*- and *W*-optimal stable matchings. Further, it is also known that when preferences are strict the set of stable matchings is a complete lattice. We give an alternative proof to the above facts through our formulation.

3. Formulation

We provide an alternative formulation to address marriage problems. To do so we introduce some notation.

Definition 3.1. A pair of functions $v \equiv (v_M, v_W)$ is called a *pre-matching* if $v_M : M \to W \cup M$ and $v_W : W \to M \cup W$ such that if $v_M(m) \neq m$ then $v_M(m) \in W$ and if $v_W(w) \neq w$ then $v_W(w) \in M$.

Let V_M and V_W denote the set of all such functions v_M and v_W, respectively. Let $V := V_M \times V_W$ denote the set of all pre-matchings v. We can think of V_M as the set of vectors in $\times_{m \in M}(W \cup \{m\})$ and of V_W as that in $\times_{w \in W}(M \cup \{w\})$. Pre-matchings are more convenient for our formulation than matchings, and they have close relationships.

Definition 3.2.

(i) For a given matching μ, a function $v \equiv (v_M, v_W)$ defined by $v_M(m) := \mu(m)$ and $v_W(w) := \mu(w)$ for all $m \in M$ and $w \in W$ is called a pre-matching v *defined* by a matching μ.
(ii) We say a pre-matching v *induces* a matching μ if a function μ defined by $\mu(m) := v_M(m)$ and $\mu(w) := v_W(w)$ is a matching.
(iii) A matching μ and a pre-matching v are said to be *equivalent* if μ defines v and v induces μ.

Note that every matching defines an equivalent pre-matching while a pre-matching may fail to induce a matching. In Example 2.5, matching μ' defines a pre-matching $v' \equiv (v'_M, v'_W)$ such that $v'_M(m_1) = w_1$, $v'_M(m_2) = w_3$, $v'_M(m_3) = w_2$, $v'_W(w_1) = m_1$, $v'_W(w_2) = m_3$ and $v'_W(w_3) = m_2$. On the other hand, a pre-matching v such that $v_M(m_1) = w_1$, $v_M(m_2) = w_3$, $v_M(m_3) = w_2$, $v_W(w_1) = m_3$, $v_W(w_2) = m_1$ and $_W(w_3) = m_3$ does not induce a matching because, for example, $v_M(m_1) = w_1$ but $v_W(w_1) = m_3 \neq m_1$. This leads to the following observation, which will be useful later.

Remark 3.3. *A pre-matching v induces a matching iff v is such that $v_M(m) = w$ iff $m = v_W(w)$. Therefore, if v induces a matching and $v_M(m) = w$ (or, equivalently, $v_W(w) = m$), then $v_W \circ v_M(m) = m$ and $v_M \circ v_W(w) = w$.*

We will show that when preferences are strict the set of stable matchings is identified with the set of solutions $v \equiv (v_M, v_W)$ to the set of equations (3.3a,b) below after an appropriate translation between matchings and pre-matchings.

To derive such equations, suppose a matching μ is stable and let v be the pre-matching defined by μ. Then from the definition of stability and Lemma 2.4, v satisfies

$$v_M(m) \geq_m m \text{ and } v_W(w) \geq_w w \text{ for } \forall m \in M, \forall w \in W, \tag{3.1}$$

$$\text{and } \nexists(m,w) \text{ such that } \{w >_m v_M(m) \text{ and } m \geq_w v_W(w)\} \text{ or } \{m >_w v_W(w) \text{ and } w \geq_m v_M(m)\}. \tag{3.2}$$

It is immediate that when preferences are strict, these conditions are equivalent to v being a solution to the following set of equations:[1]

$$v_M(m) = \max_{>_m}\{w \in W | m \geq_w v_W(w)\} \cup \{m\}, \ \ \forall m \in M, \tag{3.3a}$$

$$v_W(w) = \max_{>_w}\{m \in M | w \geq_m v_M(m)\} \cup \{w\}, \ \ \forall w \in W . \tag{3.3b}$$

In Eq. (3.3a) maximization is taken with respect to each man m's preference ordering $>_m$ over the set $W \cup \{m\}$ under the constraint $m \geq_w v_W(w)$. Since there is a finite number of agents and we have assumed strict preferences, the RHS of (3.3a,b) is well defined and singleton for each m and each w. Therefore, we have proved part (i) of the next proposition.

Proposition 3.4. *Let the assumption of strict preferences hold. Then*

  (*i*) *If a matching μ is stable, then the pre-matching v defined by μ solves Eqs. (3.3a,b).*
  (*ii*) *If a pre-matching v solves Eqs. (3.3a,b), then v induces a matching μ, which is stable.*

To prove part (ii) of the proposition, we use the following lemma:

Lemma 3.5. *Let the assumption of strict preferences hold. If $v = (v_M, v_W)$ solves Eqs. (3.3a,b), then the following conditions are equivalent*:

  (*i*) $w \geq_m v_M(m)$ *and* $m \geq_w v_W(w)$,
  (*ii*) $w = v_M(m)$ *and* $m = v_W(w)$,
  (*iii*) $w = v_M(m)$,
  (*iv*) $m = v_W(w)$.

Proof. (ii⇒i), (ii⇒iii), and (ii⇒iv) are immediate. (i⇒ii): Assume (i). Suppose $w >_m v_M(m)$ or $m >_w v_W(w)$. Either case contradicts that $v = (v_M, v_W)$ solves Eqs. (3.3a,b). So it must be $w =_m v_M(m)$ and $m =_w v_W(w)$. But, with the assumption of strict preferences, this implies that $w = v_M(m)$ and $m = v_W(w)$. (iii⇒iv⇒ii): Assume $w = v_M(m)$. Suppose $m >_w v_W(w)$. But this contradicts (3.3b).

---

[1]Roth et al. (1993) characterize stable matchings as solutions to a linear programming problem. A referee pointed out a similarity between inequalities used in our Eqs. (3.3a,b) and those in their linear programming formulation (their individual rationality constraints (6) and stability constraints (7)).

Suppose $m <_w v_W(w)$. This contradicts (3.3a). So it must be $m =_w v_W(w)$, which implies $m = v_W(w)$ under the assumption of strict preferences. $\square$

Proof of Proposition 3.4, part (ii). Let the assumption of strict preferences hold. Suppose v satisfies Eqs. (3.3a,b). From Remark 3.3 and parts (iii) and (iv) of Lemma 3.5, v induces some matching μ. Also, v satisfies conditions (3.1) and (3.2). It means that μ satisfies conditions (IR) and (S′), which implies conditions (IR) and (S). $\square$

Next we will show that the set of solutions to Eqs. (3.3a,b) is non-empty and is a complete lattice. First we define partial orderings on sets V_M, V_W, and V.

Definition 3.6. Let $v \equiv (v_M, v_W) \in V$.

(i) Define a partial ordering $\geq_M$ on V_M by $v_M \geq_M v'_M$ iff $v_M(m) \geq_m v'_M(m)$ for all $m \in M$.
(ii) Define a partial ordering $\geq_W$ on V_W by $v_W \geq_W v'_W$ iff $v_W(w) \geq_w v'_W(w)$ for all $w \in W$.
(iii) Define a partial ordering $\geqq_M$ on V by $v \geqq_M v'$ iff $v_M \geq_M v'_M$ and $v_W \leq_W v'_W$.

Consider a mapping $T \equiv (T_1, T_2)$, where $T_1 : V \rightarrow V_M$ and $T_2 : V \rightarrow V_W$ are defined by the right-hand sides of Eqs. (3.3a) and (3.3b), respectively. We can prove the following proposition by applying Tarski's fixed point theorem (Tarski, 1955).

Proposition 3.7. *The set V^* of solutions to Eqs. (3.3a,b) is non-empty and $\langle V^*, \geqq_M \rangle$ is a complete lattice.*

Proof. We need to show that the set of fixed points of mapping T has the above property. Note that $\langle V, \geqq_M \rangle$ is a finite lattice and thus complete. Apparently, $T : V \rightarrow V$. We need to show that mapping T is non-decreasing in v with respect to $\geqq_M$. Consider any $v' = (v'_M, v'_W)$ and $v'' = (v''_M, v''_W)$ such that $v'' \geqq_M v'$ (i.e., $v''_M \geq_M v'_M$ and $v''_W \leq_W v'_W$). Then

$$
\begin{aligned}
T_1(v'')(m) &=_m \max_{>_m} \{w \in W | m \geq_w v''_W(w)\} \cup \{m\} \\
&\geq_m \max_{>_m} \{w \in W | m \geq_w v'_W(w)\} \cup \{m\} \\
&=_m T_1(v')(m).
\end{aligned}
$$

The above inequality $\geq_m$ just follows from the fact that $\{w \in W : m \geq_w v''_W(w)\} \supseteq \{w \in W : m \geq_w v'_W(w)\}$. Thus, $T_1 v'' \geq_M T_1 v'$. Similarly, $T_2 v'' \leq_W T_2 v'$. Hence, $T v'' \geqq_M T v'$. $\square$

When preferences are strict, Proposition 3.4 and Proposition 3.7 imply that we can identify the set of stable matchings with the set V^* of solutions to Eqs. (3.3a,b). In such cases, we sometimes call $v \in V^*$ itself a stable matching, and call V^* the set of stable matchings. Let $\bar{v} \equiv (\bar{v}_M, \underline{v}_W)$ and $\underline{v} \equiv (\underline{v}_M, \bar{v}_W)$ denote, respectively, the greatest and smallest elements (with respect to $\geqq_M$) in V^*. A matching $\bar{v}$ is an *M*-optimal stable matching and a matching $\underline{v}$ a *W*-optimal stable matching; every man likes and every woman dislikes $\bar{v}$ at least as much as any other stable matching, and the opposite is true for $\underline{v}$. The *M*- and *W*-optimal equilibria, $\bar{v}$ and $\underline{v}$, can be found by a simple iterative procedure: to obtain $\bar{v}$, set $\bar{v}^0 \equiv (\bar{v}^0_M, \underline{v}^0_W)$ such that $\bar{v}^0_M(m) := \max_{>_m} w$ s.t. $w \in W \cup \{m\}$ for all m and $\underline{v}^0_W(w) := w$

for all w. Define a sequence $\bar{v}^n \equiv (\bar{v}_M^n, \underline{v}_W^n)$ by $\bar{v}^n := T\bar{v}^{n-1}$ for $n \geq 1$. Since the sets M and W are finite, the sequence $\bar{v}^n$ converges to $\bar{v} := \lim \bar{v}^n$ after a finite n. The limit $\bar{v}$ is the M-optimal stable matching. Similarly, to obtain the W-optimal equilibrium, define a sequence $\underline{v}^n \equiv (\underline{v}_M^n, \bar{v}_W^n)$ by $\underline{v}^n := T\underline{v}^{n-1}$ starting with $\underline{v}_M^0(m) := m$ and $\bar{v}_W^0(w) := \max_{>_w} m$ s.t. $m \in M \cup \{m\}$. Then, $\underline{v} := \lim \underline{v}^n$ is the W-optimal stable matching.

4. When preferences are not strict

When preferences are not strict, the set of solutions to Eqs. (3.3a,b) do not fully characterize the set of stable matchings as it does in Proposition 3.4. The problem is that when preferences are not strict, a solution v to Eqs. (3.3a,b) may fail to induce a matching, and a pre-matching defined by some stable matching may not satisfy Eqs. (3.3a,b). However, we can still show the existence of a stable matching using these equations by introducing tie-breaking rules into not strict preferences.

Note that the method in Section 3 would work if the preferences were strict. So let us transform a not-necessarily-strict preference ordering of each agent into a strict ordering by arbitrarily assigning strict orders among the group of alternatives the agent regards as equivalent, keeping orders for the rest of the alternatives as in the original preference ordering. Then, there is a solution v to Eqs. (3.3a,b) with respect to the transformed preferences. And v induces a matching, which is stable under the original preferences. Therefore, this reproves the result by Gale and Shapley (1962) that every marriage problem has a stable matching, without reference to the deferred acceptance algorithm.

5. Discussion

An interesting point of our formulation is that Eqs. (3.3a,b) look as if they were defining a sort of Nash equilibrium in some non-cooperative game. In Adachi (1998), the author uses these equations to show that the set of equilibrium outcomes in a search theoretic marriage problem with negligible search costs reduces to the set of stable matchings in a corresponding Gale–Shapley marriage problem.

This new characterization of stable matchings with strict preferences is very prone to comparative statics analyses as it relies on monotonicity (see Milgrom and Roberts, 1994; Milgrom and Shannon, 1994). It can give a simple proof to one of the well-known comparative statics results that when men extend their lists of acceptable women, the men are not better off and the women are not worse off (see, for example, Theorem 2.24 in Roth and Sotomayor, 1990).

A referee suggested extending this formulation to the college admissions problem. However, I am unable to provide another proof to the strong lattice properties of the stable matchings in the college admissions problem (see Section 5.6 of Roth and Sotomayor, 1990) within our formulation, independently of their Lemma 5.25. The difficulty lies in treating colleges' preferences over groups of students, not just preferences over individual students, in our formulation.

One attractive question that remains to be answered is whether a similar characterization is possible for the assignment games studied by Shapley and Shubik (1972). Using a Nash bargaining solution, Rochford (1984) considers a rebargaining process between players in Shapley–Shubik assignment

games and identifies some interior points in the core as stationary points of the rebargaining process. Following her work, Roth and Sotomayor (1988) observe a certain monotonicity in the rebargaining process and show that these interior points have a lattice property in (an extension of) the assignment games. It might be possible to identify the core itself as the stationary points of some rebargaining process by including a larger class of bargaining solutions.

Acknowledgements

References

Adachi, H., 1998. A search-theoretic approach to two-sided matching. Ph.D. dissertation, SUNY-Buffalo.

Gale, D., Shapley, L., 1962. College admissions and the stability of marriage. American Mathematical Monthly 69, 9–15.

Milgrom, P., Roberts, J., 1994. Comparing equilibria. American Economic Review 84, 441–459.

Milgrom, P., Shannon, C., 1994. Monotone comparative statics. Econometrica 62, 157–180.

Rochford, S., 1984. Symmetrically pairwise-bargained allocations in an assignment market. Journal of Economic Theory 34, 262–281.

Roth, A., Rothblum, U., Vande Vate, J., 1993. Stable matchings, optimal assignments, and linear programming. Mathematics of Operations Research 18, 803–828.

Roth, A., Sotomayor, M., 1988. Interior points in the core of two-sided matching markets. Journal of Economic Theory 45, 85–101.

Roth, A., Sotomayor, M., 1990. Two-sided Matching, Cambridge University Press, Cambridge.

Shapley, L., Shubik, M., 1972. The assignment game I: the core. International Journal of Game Theory 1, 23–28.

Tarski, A., 1955. A lattice-theoretic fixpoint theorem and its applications. Pacific Journal of Mathematics 5, 285–309.