Mathematical Programming 54 (1992) 57-67
North-Holland

Characterization of stable matchings as extreme points of a polytope

Uriel G. Rothblum

Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Haifa 32000, Israel, and RUTCOR - Rutgers Center for Operations Research, Hill Center for Mathematical Sciences, Rutgers University, New Brunswick, NJ 08904, USA

Received 17 April 1989
Revised manuscript received 18 January 1990; 14 May 1990

The purpose of this paper is to extend a modified version of a recent result of Vande Vate (1989) which characterizes stable matchings as the extreme points of a certain polytope. Our proofs are simpler and more transparent than those of Vande Vate.

Key words: Matchings, stability, extreme points, polytope.

1. Introduction

Gale and Shapley (1962) considered a model where n men and n women are to be matched where each individual has a (strict) preference relation over the group of individuals of the opposite sex. They defined a matching to be *stable* if there exist no man and woman who prefer each other over their corresponding mates. Gale and Shapley then showed that a stable matching exists for each profile of individual preferences. Their constructive proof relies on an algorithm which has an interesting interpretation in the context of courting and the mechanism of mate-selection.

Much research on the above matching problem and related modifications has followed the original work of Gale and Shapley. The books of Knuth (1976), Irwing and Gusfield (1989) and Roth and Sotomayor (1990) can be viewed as landmarks in the development; in particular, the latter two contain extended lists of references to contributions that have taken place in the last three decades. An interesting study of the labor market for medical interns and its relation to the matching problem is given in Roth (1984).

Evidently, each matching of n men to n women can be represented by an assignment matrix. Vande Vate (1989) formulated a system of linear inequalities such that the set of assignment matrices which satisfy it are precisely those that correspond to stable matchings. He also obtained a surprising characterization of these assignment matrices as the extreme points of a corresponding polytope.

In the current paper we resolve a problem raised by Vande Vate by obtaining an extension of his results to situations where singlehood is permitted and individuals are allowed to stay unmatched. The extension requires a modification of the system of linear inequalities used by Vande Vate and the proofs we provide turn out to be simpler and more transparent.

We present our main results in Section 2 and discuss their specialization to the original Gale–Shapley model in Section 3. Finally, in Section 4, we consider the problem of many-to-one matchings.

2. Characterization of stable matchings

Formally, let M and W denote two finite disjoint sets. These sets will have the interpretation of groups of *men* and *women*, respectively. For each $m \in M$ we have a subset W_m of W and a strict preference relation $>_m$ that is defined on W_m. Similarly, for each $w \in W$ we have a subset M_w of M and a strict preference relation $>_w$ on M_w. For $k \in M \cup W$, we will use the standard notation $<_k$ to denote the derived reverse relation of $>_k$. Also, we will use the notation $\geqslant_k$ and $\leqslant_k$ to denote the relations "$>_k$ or $=$" and "$<_k$ or $=$", respectively. For $m \in M$ we say that the members of W_m are *acceptable* to m and that those in $W \setminus W_m$ are *unacceptable* to him. Similarly, for $w \in W$, we say that the members of M_w are *acceptable* to w and that those in $M \setminus M_w$ are *unacceptable* to her. Also, the set of *acceptable pairs*, denoted by A, is the set $A \equiv \{(m, w)\colon m \in M_w \text{ and } w \in W_m\}$. This model modifies the original Gale–Shapley model as individuals are allowed to prefer staying single over being matched to some of the members of the opposite sex.

Each matching of men to women can be represented by a (*Partial*) *assignment matrix*, i.e., a matrix $x = \{x_{ij}\}_{i \in M, j \in W}$ whose elements are all integers and which satisfies

$$\sum_{j \in W} x_{ij} \leqslant 1 \quad \text{for all } i \in M, \tag{1}$$

$$\sum_{i \in M} x_{ij} \leqslant 1 \quad \text{for all } j \in W, \tag{2}$$

and

$$x_{ij} \geqslant 0 \quad \text{for all } i \in M \text{ and } j \in W. \tag{3}$$

Of course, all elements of such a matrix are either 1 or 0. In the remainder of this paper we identify matchings with the corresponding assignment matrices. In particular, given such a matrix x, we say that $m \in M$ and $w \in W$ are *matched to each other* if $x_{mw} = 1$, in which case we write $w = w(x, m)$ and $m = m(x, w)$. Also, we say that $k \in M \cup W$ *has a mate* if that individual is matched to someone (of the opposite sex).

A matching is defined to be *stable* if no individual is matched to an unacceptable person and if there exist no man and woman who prefer each other to the outcome of the matching. Formally, stability of a matching is expressed by the following conditions on the corresponding assignment matrix, say x: first,

$$x_{ij} = 0 \quad \text{for all } (i, j) \in (M \times W) \setminus A, \tag{4}$$

and, second, there exists no pair $(m, w) \in A$ for which either of the following four conditions holds: (i) both m and w have mates, $w(x, m) <_m w$ and $m(x, w) <_w m$; or (ii) m has a mate and $w(x, m) <_m w$, whereas w has no mate; or (iii) w has a mate and $m(x, w) <_w m$, whereas m has no mate; or (iv) both m and w have no mates.

The following lemma modifies and generalizes an observation of Vande Vate (1989) who considered the original Gale–Shapley model. It expresses stability via a set of linear inequalities.

Lemma 1. *Let x be a matching. Then x is stable if and only if x satisfies* (4) *and the following linear inequalities*

$$\sum_{j >_m w} x_{mj} + \sum_{i >_w m} x_{iw} + x_{mw} \geq 1 \quad \textit{for all } (m, w) \in A. \tag{5}$$

(*The summations in* (5) *are defined over* $\{j \in W_m : j >_m w\}$ *and* $\{i \in M_w : i >_w m\}$, *respectively, and the notations* $\sum_{j >_m w}$ *and* $\sum_{i >_w m}$ *are used for brevity.*)

Proof. We observe that (5) is violated if and only if for some $(m, w) \in A$,

$$\sum_{j >_m w} x_{mj} = \sum_{i >_w m} x_{iw} = x_{mw} = 0,$$

i.e., either m has no mate or he does have a mate and $w(x, m) <_m w$, and in addition, either w has no mate or she does have a mate $m(x, w) <_w m$. We get four pairs of conditions which are precisely the conditions that characterize stability. So, stability is equivalent to the combination of conditions (4) and (5). □

Lemma 1 shows that $x = \{x_{ij}\}_{i \in M, j \in W}$ represents a stable matching if and only if x is an integer solution of (1)–(5). We next consider solutions to (1)–(5) that are not necessarily integral. Let $x = \{x_{ij}\}_{i \in M, j \in W}$ satisfy (1)–(5). Define $S_M(x) \equiv \{m \in M : \sum_{j \in W} x_{mj} > 0\}$ and $S_W(x) \equiv \{w \in W : \sum_{i \in M} x_{iw} > 0\}$. Further, for $m \in S_M(x)$, let $w^*(x, m)$ and $w_*(x, m)$ be, respectively, the most preferred and least preferred elements in $\{j \in W : x_{mj} > 0\}$ with respect to the preference relation $>_m$ (of course, (4) assures that $\{j \in W : x_{mj} > 0\} \subseteq W_m$, so $>_m$ is defined on this set). Also, for $w \in S_W(x)$ we define $m^*(x, w)$ and $m_*(x, w)$ correspondingly. We observe that x is integer if and only if for all $m \in S_M(x)$, $\sum_{j \in W} x_{mj} = 1$ and $w^*(x, m) = w_*(x, m)$, or equivalently, if and only if for all $w \in S_W(x)$, $\sum_{i \in M} x_{iw} = 1$ and $m^*(x, w) = m_*(x, w)$.

The following two lemmas provide useful tools for studying solutions of (1)–(5) which are not necessarily integral.

Lemma 2. *Let* $x=\{x_{ij}\}_{i\in M, j\in W}$ *satisfy* (1)–(5) *and let* $(m, w)\in A$. *Then*

$$\{[m\notin S_M(x)] \text{ or } [m\in S_M(x) \text{ and } w \geqslant_m w^*(x,m)]\}$$

$$\Rightarrow \left[\sum_{i\in M} x_{iw}=1 \text{ and } m\leqslant_w m_*(x,w)\right], \tag{6}$$

and

$$[m\in S_M(x) \text{ and } w=w^*(x,m)] \Leftrightarrow \left[\sum_{i\in M} x_{iw}=1 \text{ and } m=m_*(x,w)\right]. \tag{7}$$

Further, if

$$\sum_{j>_m w} x_{mj}+\sum_{i>_w m} x_{iw}+x_{mw}=1, \tag{8}$$

then the reverse implication of (6) *holds as well. Finally,* $m\in S_M(x)$ *if and only if* $\sum_{j\in W} x_{mj}=1$.

Proof. We start with (6). Assume that $m\notin S_M(x)$ or that $m\in S_M(x)$ and that $w\geqslant_m w^*(x,m)$. Then $\sum_{j>_m w} x_{mj}=0$, and therefore, by (5), $\sum_{i>_w m} x_{iw}+x_{mw}\geqslant 1$. So, $1\leqslant\sum_{i>_w m} x_{iw}+x_{mw}\leqslant\sum_{i\in M} x_{iw}\leqslant 1$, implying that $\sum_{i\in M} x_{iw}=1$ and $x_{iw}=0$ for all $i\in M_w$ with $i<_w m$. In particular, $w\in S_W(x)$ and $m\leqslant_w m_*(x,w)$.

Next, to see the implication $\Rightarrow$ of (7), assume that $m\in S_M(x)$ and $w=w^*(x,m)$. Then $x_{mw}>0$, implying that $w\in S_W(x)$ and $m\geqslant_w m_*(x,w)$. Also, by the (established) implication (6), $\sum_{i\in M} x_{iw}=1$ and $m\leqslant_w m_*(x,w)$. Hence, the right-hand side assertion of (7) holds, completing the proof of the implication $\Rightarrow$ of (7). It follows that the map $w^*(x,\cdot)$ is one-to-one from $S_M(x)$ into $F_W(x)\equiv\{w\in W: \sum_{i\in M} x_{iw}=1\}$ (it is one-to-one because if $w=w^*(x,m_1)=w^*(x,m_2)$ for m_1 and m_2 in $S_M(x)$, then $w\in F_W(x)\supseteq S_W(x)$ and $m_1=m_*(x,w)=m_2$). Letting $|\cdot|$ denote cardinality we have that

$$|F_W(x)|=\sum_{j\in F_W(x)}\sum_{i\in M} x_{ij}=\sum_{i\in S_M(x)}\sum_{j\in F_W(x)} x_{ij}\leqslant\sum_{i\in S_M(x)} 1=|S_M(x)|,$$

implying that the map $w^*(x,\cdot)$ maps $S_M(x)$ onto $F_W(x)$. In particular, we conclude that $|F_W(x)|=|S_M(x)|$ and that for each $i\in S_M(x)$, $\sum_{j\in F_W(x)} x_{ij}=1$, implying the last conclusion of our lemma. Next, to see the implication $\Leftarrow$ of (7), assume that $\sum_{i\in M} X_{iw}=1$ and $m=m_*(x,w)$. As $w^*(x,\cdot)$ is onto, there is a $m'\in S_M(x)$ satisfying $w=w^*(x,m')$. Then by the established implication $\Rightarrow$ of (7), $m'=m_*(x,w)=m$. So, $m=m'\in S_M(x)$ and $w=w^*(x,m')=w^*(x,m)$, completing the proof of (7).

Finally assume that (8) holds, that $\sum_{i\in M} x_{iw}=1$ (assuring that $w\in S_W(x)$) and that $m\leqslant_w m_*(x,w)$. Then

$$\sum_{j>_m w} x_{mj}+\sum_{i>_w m,} x_{iw}+x_{mw}=1=\sum_{i\in M} x_{iw}=\sum_{i>_w m} x_{iw}+x_{mw},$$

implying that $\sum_{j>_m w} x_{mj}=0$, i.e., either $m\notin S_M(x)$, or $m\in S_M(x)$ and $w^*(x,m)\leqslant_m w$. So, we have established the reverse implication of (6) under the assumption that (8) holds. $\square$

By interchanging the roles of men and women we get implications that are symmetric to those established in Lemma 2. They are summarized in the following lemma.

Lemma 3. *Let* $x = \{x_{ij}\}_{i\in M, j\in W}$ *satisfy* (1)–(5), *and let* $(m, w) \in A$. *Then*

$$\{[w \notin S_W(x)] \text{ or } [w \in S_W(x) \text{ and } m \geqslant_w m^*(x, w)]\}$$
$$\Rightarrow \left[\sum_{j\in W} x_{mj} = 1 \text{ and } w \leqslant_m w_*(x, m)\right], \tag{9}$$

and

$$[w \in S_W(x) \text{ and } m = m^*(x, w)] \Leftrightarrow \left[\sum_{j\in W} x_{mj} = 1 \text{ and } w = w_*(x, m)\right]. \tag{10}$$

Further, if (8) *holds, then the reverse implication of* (9) *holds as well. Finally,* $m \in S_M(X)$ *if and only if* $\sum_{i\in M} x_{iw} = 1$. □

Remark. In view of the last conclusion in Lemmas 2 and 3, one can replace the conditions "$\sum_{i\in M} x_{iw} = 1$" and "$\sum_{j\in W} x_{mj} = 1$" in (6), (7), (9) and (10) by the conditions "$w \in S_W(x)$" and "$m \in S_M(x)$", respectively.

Lemmas 2 and 3 demonstrate that men and women have conflicting interests. This phenomenon has been previously captured in the context of stable matching in the Decomposition Lemma of Knuth (1976); see also Roth and Sotomayor (1990, Section 2.5). Before continuing with the development of our main result, we demonstrate the useful ness of Lemmas 2 and 3 by showing that they provide an alternative proof for the Decomposition Lemma.

Lemma 4 (Knuth, 1976). *Let x and y be two stable matchings. Consider*

$$K(x) \equiv \{k \in M \cup W\text{: } k \text{ prefers the outcome under } x \text{ to the outcome under } y\}$$

and

$$K(y) \equiv \{k \in M \cup W\text{: } k \text{ prefers the outcome under } y \text{ to the outcome under } x\}.$$

Then $w(x, \cdot)$ *maps* $M \cap K(x)$ *onto* $W \cap K(y)$ *and* $M \cap K(y)$ *onto* $W \cap K(x)$; *further, the same conclusion holds for* $w(y, \cdot)$.

Proof. Suppose $m \in M \cap K(x)$. Let $z \equiv \frac{1}{2}x + \frac{1}{2}y$. As m prefers x to y, we have that m is matched under x, i.e., $m \in S_M(x)$. Further, as $S_M(z) \supseteq S_M(x)$ it also follows that $m \in S_M(z)$. Now, as x and y satisfy (1)–(5), so does z. Hence, Lemma 2 applies to z and we conclude from the last part of that lemma that $\sum_{j\in W} z_{mj} = 1$, implying that $m \in S_M(y)$, so m is matched under y. As the assertion that m prefers x to y means that $w \equiv w(x, m) = w^*(y, m) >_m w(y, m)$, Lemma 2 implies that $w \in S_W(y)$ and that $m(x, w) = m <_w m_*(y, w) = m(y, w)$; so, $w \in K(y)$. By interchanging the roles of m and w and applying Lemma 3 to x we also get that if $w \in K(y)$ then $m(x, w) \in K(x)$. Thus, we have shown that $w(x, \cdot)$ maps $M \cap K(x)$ onto $W \cap K(y)$. The remaining conclusions of the lemma follow by interchanging the roles of x and y and of men and women. □

Recall that an *extreme point* of a convex set C is a point x in C for which there exist no x' and x'' in C, both distinct from x, and $0 < \alpha < 1$, such that $x = (1-\alpha)x' + \alpha x''$. We are now ready for the characterization of stable matchings, i.e., the integer solutions of (1)-(5). Vande Vate (1989) considered matching problems in which singlehood is not permitted and used a variant of the linear inequality system (1)-(5).

Theorem 1. *Let C be the set of solutions of* (1)-(5). *Then the integer points in C are precisely its extreme points.*

Proof. Assume that x is an integer solution of (1)-(5). As integer solutions of the system consisting of (1)-(3) are known to be extreme points of the set defined by these constraints (Birkhoff's Theorem), we conclude that for no $x' \neq x$ and $x'' \neq x$ satisfying (1)-(3) and $0 < \alpha < 1$, $x = (1-\alpha)x' + \alpha x''$. In particular, no such representation is possible with x' and x'' that satisfy (1)-(5). So, x is an extreme point of C.

Next, let x be an extreme point of C. In particular, x satisfies (1)-(5). Define the matrices $z^* = \{(z^*)_{mw}\}_{m \in M, w \in W}$, $z_* = \{(z_*)_{mw}\}_{m \in M, w \in W}$ and $z = \{z_{mw}\}_{m \in M, w \in W}$ by

$$(z^*)_{mw} \equiv \begin{cases} 1 & \text{if } m \in S_M(x) \text{ and } w = w^*(x, m), \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

$$(z_*)_{mw} \equiv \begin{cases} 1 & \text{if } m \in S_M(x) \text{ and } w = w_*(x, m), \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

and

$$(z)_{mw} \equiv (z^*)_{mw} - (z_*)_{mw}. \tag{13}$$

We observe that Lemmas 2 and 3 and the remark following these lemmas show that

$$(z^*)_{mw} \equiv \begin{cases} 1 & \text{if } w \in S_W(x) \text{ and } m = m_*(x, w), \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

and

$$(z_*)_{mw} \equiv \begin{cases} 1 & \text{if } w \in S_W(x) \text{ and } m = m^*(x, w), \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

We will next demonstrate that

$$\left[\sum_{j \in W} x_{ij} = 1\right] \Rightarrow \left[\sum_{j \in W} z_{ij} = 0\right], \quad i \in M, \tag{16}$$

$$\left[\sum_{i \in M} x_{ij} = 1\right] \Rightarrow \left[\sum_{i \in M} z_{ij} = 0\right], \quad j \in W, \tag{17}$$

$$[x_{ij} = 0] \Rightarrow [z_{ij} = 0], \quad i \in M \text{ and } j \in W, \tag{18}$$

$$z_{ij} = 0, \quad (i, j) \in (M \times W) \backslash A, \tag{19}$$

and

$$\left[\sum_{j>_m w} x_{mj} + \sum_{i>_w m} x_{iw} + x_{mw} = 1\right] \Rightarrow \left[\sum_{j>_m w} z_{mj} + \sum_{i>_w m} z_{iw} + z_{mw} = 0\right] \quad (m, w) \in A. \tag{20}$$

It will then follow that for sufficiently small positive ε, $x+\varepsilon z$ and $x-\varepsilon z$ satisfy (1)-(5). As $x=\frac{1}{2}(x+\varepsilon z)+\frac{1}{2}(x-\varepsilon z)$, the extremality of x will imply that $x+\varepsilon z = x-\varepsilon z = x$, i.e., $z=0$, or equivalently, $z^* = z_*$. We will then conclude that for all $m \in S_M(x)$, $w^*(x, m) = w_*(x, m)$ and the integrality of x will follow.

So, it remains to establish (16)-(20). We first observe that the definitions of z^* and z_* given in (11) and (12) assure that if $m \in S_M(x)$, then $\sum_{j\in W}(z^*)_{mj} = \sum_{j\in W}(z_*)_{mj} = 1$, implying that $\sum_{j\in W} z_{mj} = 0$. So, (16) holds. Next, (17) follows from similar arguments using the representations of z^* and z_* given in (14) and (15). Also, (18) is straightforward as $x_{ij}=0$ assures that if $i \in S_M(x)$ then $j \notin \{w^*(x, i), w_*(x, i)\}$. Next, (19) is immediate from (18) and (4). Finally, to establish (20), assume that $(m, w) \in A$ and that

$$\sum_{j>_m w} x_{mj} + \sum_{i>_w m} x_{iw} + x_{mw} = 1, \tag{21}$$

and we will show that

$$\sum_{j>_m w} z_{mj} + \sum_{i>_w m} z_{iw} + z_{mw} = 0. \tag{22}$$

We establish (22) by considering three cases. First assume that either $m \notin S_M(x)$ or that $m \in S_M(x)$ with $w \geqslant_m w^*(x, m)$. Then the definition of z^* and z_* via (11) and (12) shows that

$$\sum_{j>_m w} (z^*)_{mj} = \sum_{j>_m w} (z_*)_{mj} = 0. \tag{23}$$

Further, we conclude from Lemma 2 that $w \in S_W(x)$ and $m \leqslant_w m_*(x, w) \leqslant_w m^*(x, w)$. Hence, (14) and (15) imply that

$$\sum_{i>_w m} (z^*)_{iw} + (z^*)_{mw} = \sum_{i>_w m} (z_*)_{iw} + (z_*)_{mw} = 1, \tag{24}$$

and (22) follows directly from (23)-(24). Similar arguments show that (21) implies (22) when either $w \notin S_W(x)$ or $w \in S_W(x)$ with $m \geqslant_w m^*(x, w)$. It remains to consider the case where $m \in S_M(x)$, $w \in S_W(x)$, $w <_m w^*(x, m)$ and $m <_m m^*(x, w)$. As $m \in S_M(x)$ and $w <_m w^*(x, m)$, the definition of z^* in (11) implies that $\sum_{j>_m w}(z^*)_{mj} = 1$ and $(z^*)_{mw} = 0$. Also, by Lemma 3, (21) implies the reverse implication of (6) and the remark following Lemma 3 shows that the assertion "$\sum_{j\in W} x_{mj} = 1$" can be replaced by the assertion that "$m \in S_M(x)$". As $m \in S_M(x)$, $w \in S_W(x)$ and $w <_m w^*(x, m)$, we conclude from the (modified) reverse implication of (6) that $m >_w m_*(x, w)$. It now follows directly from (14) that, $\sum_{i>_w m}(z^*)_{iw} = 0$. So, we have that

$$\sum_{j>_m w} (z^*)_{mj} = 1, \qquad (z^*)_{mw} = 0 \quad \text{and} \quad \sum_{i>_w m} (z^*)_{iw} = 0, \tag{25}$$

and therefore,

$$\sum_{j>_m w} (z^*)_{mj} + \sum_{i>_w m} (z^*)_{iw} + (z^*)_{mw} = 1. \tag{26}$$

Symmetric arguments show that $\sum_{i>_w m} (z_*)_{iw} = 1$, $(z_*)_{mw} = 0$ and $\sum_{j>_m w} (z_*)_{mj} = 0$. So,

$$\sum_{i>_w m} (z_*)_{iw} + \sum_{j>_m w} (z_*)_{mj} + (z_*)_{mw} = 1 \tag{27}$$

and (26) and (27) establish (22) in the third and final case. □

3. The orginal Gale–Shapley model

We next show how our results specialize to the original Gale–Shapley model in which preferences of individuals are defined over the set of all members of the opposite sex and singlehood is not permitted. In particular, we have that $M_w = M$ for all $m \in M$, $W_m = W$ for all $w \in W$ and a matching corresponds to a *complete assignment matrix*, i.e., a matrix $x = \{x_{ij}\}_{i \in M, j \in W}$ with integer coordinates which satisfies

$$\sum_{j \in W} x_{ij} = 1 \quad \text{for all } i \in M, \tag{1'}$$

$$\sum_{i \in M} x_{ij} = 1 \quad \text{for all } j \in W, \tag{2'}$$

and (3). We note that (1′) and (2′) modify (1) and (2) respectively, by replacing the corresponding inequalities by equalities. Also, existence of a vector x satisfying (1′) and (2′) is equivalent to the assertion that the number of men and the number of women coincide. The next lemma shows that in the above case (1)–(5) are equivalent to the combination of (1′), (2′), (3) and (5).

Lemma 5. *Assume that the number of elements in M and W coincide, that $W_m = W$ for all $m \in M$ and that $M_w = M$ for all $w \in W$. Then $x = \{x_{ij}\}_{i \in M, j \in W}$ satisfies* (1)–(5) *if and only if x satisfies* (1′), (2′), (3) *and* (5). *In particular, in this case $M = S_M(x)$ and $W = S_W(x)$.*

Proof. Since (1′) implies (1), (2′) implies (2) and (4) is trivial under the assumption that $W_m = W$ for all $m \in M$ and that $M_w = M$ for all $w \in W$, it suffices to show that under our assumptions, (1)–(5) imply (1′) and (2′). So, assume that x satisfies (1)–(5). Letting $|\cdot|$ denote cardinality, we have from the last conclusion of Lemmas 2 and 3 that $|S_M(x)| = \sum_{j \in W} \sum_{i \in M} x_{ij} = |S_W(x)|$. Hence, as $|M| = |W|$, we conclude that $|M \setminus S_M(x)| = |W \setminus S_W(x)|$. Now, if $M \setminus S_M(x)$ and $W \setminus S_W(x)$ are non-empty, we get a violation of (5). So, it follows that $M = S_M(x)$ and $W = S_W(x)$ and the last conclusion in Lemmas 2 and 3 imply (1′) and (2′). □

Lemma 5 allows us to simplify the statements of Lemmas 1-3 and Theorem 1 when the original Gale-Shapley model is considered.

Lemma 1′. *Assume that the number of elements in M and W coincide, that $W_m = W$ for all $m \in M$ and that $M_w = M$ for all $w \in W$. Let x be a complete matching. Then x is stable if and only if x satisfies* (5). □

Lemma 2′. *Assume that the number of elements in M and W coincide, that $W_m = W$ for all $m \in M$ and that $M_w = M$ for all $w \in W$. Let $x = \{x_{ij}\}_{i \in M, j \in W}$ satisfy* (1′), (2′), (3) *and* (5) *and let $m \in M$ and $w \in W$. Then*

$$[w \geqslant_m w^*(x, m)] \Rightarrow [m \leqslant_w m_*(x, w)], \tag{6′}$$

and

$$[w = w^*(x, m)] \Leftrightarrow [m = m_*(x, w)]. \tag{7′}$$

Further, if (8) *holds, then the reverse implication of* (6′) *holds as well.* □

Lemma 3′. *Assume that the number of elements in M and W coincide, that $W_m = W$ for all $m \in M$ and that $M_w = M$ for all $w \in W$. Let $x = \{x_{ij}\}_{i \in M, j \in W}$ satisfy* (1′), (2′), (3) *and* (5) *and let $m \in M$ and $w \in W$. Then*

$$[M \geqslant_w m^*(x, w)] \Rightarrow [w \leqslant_m w_*(x, m)] \tag{9′}$$

and

$$[m = m^*(x, w)] \Leftrightarrow [w = w_*(x, m)]. \tag{10′}$$

Further, if (8) *holds, then the reverse implication of* (9′) *holds as well.* □

Theorem 1′ (Vande Vate, 1988). *Assume that the number of elements in M and W coincide, that $W_m = W$ for all $m \in M$ and that $M_w = M$ for all $w \in W$. Let C' be the set of solutions of* (1′), (2′), (3) *and* (5). *Then the integer points in C' are precisely its extreme points.* □

Vande Vate (1989) obtained Theorem 1′ with (5) substituted by its modified version

$$\sum_{j >_m w} x_{mj} + \sum_{i >_w m} x_{iw} + x_{mw} \leqslant 1 \quad \text{for all } (m, w) \in A. \tag{5′}$$

Obviously, (5) and (5′) are equivalent whenever (1′) and (2′) hold; so, by Lemma 5, Theorem 1′ is equivalent to Vande Vate's result. However, (5′) cannot replace (5) in Theorem 1 itself.

We note that one can use our arguments to obtain direct proofs of the results that concern the original Gale-Shapley model, and such proofs will be somehow simpler than those given in Section 2 for the general model. The more general results of Section 2 can then be derived by embedding the model in a larger one where (strict) preferences are complete over members of the opposite sex and slack individuals (representing singlehoods). It is then possible to show that there is a one-to-one correspondence between the corresponding polytopes which preserves integrality and maps extreme points onto extreme points. We use a direct approach to study the general problem because it is less cumbersome, provides more explicit representation of the corresponding stable matchings and avoids the need to consider a lot of irrelevant preferences.

4. Extension to many-to-one matchings

We finally consider the extension of the matching problem where one of the two sexes, say the men, are allowed to be matched to more than one individual of the opposite sex. This extension captures the college admission problem where colleges are matched with a number of students, while each student is matched with a single college. It also applies to labor markets where institutions need more than one worker, but each worker is assigned to a single job. We analyze the many-to-one matchings under the assumption that the preferences of members of the group that gets multiple mates are determined only by preferences over individuals that are assigned to them, rather than being dependent on the assigned set of mates. See Roth [1985] for discussion of the restrictive nature of this assumption.

We next describe the modification of the model more formally. Following standard notation, we replace the notations M and W by C and T, and refer to their elements as *colleges* and *students*, respectively. We assume that to each $c \in C$ corresponds a positive integer q_c which is the maximal *quota* of college c. Also, to each college $c \in C$ corresponds a subset T_c of T of *acceptable students* to that *college* and a strict preference $<_c$ over T_c. Similarly, to each student $t \in T$ corresponds a subset C_t of C of *acceptable colleges* and a strict preference $<_t$ over C_t. Let A be the set of *acceptable pairs*, i.e., the set $\{(c, t)\colon c \in C_t \text{ and } t \in T_c\}$. A *matching* is an assignment of students to colleges and it can be represented by an integer matrix $(x_{ct})_{c \in C, t \in T}$ satisfying

$$\sum_{t \in T} x_{ct} \leqslant q_c \quad \text{for all } c \in C, \tag{1''}$$

$$\sum_{c \in C} x_{ct} \leqslant 1 \quad \text{for all } t \in T, \tag{2''}$$

and

$$x_{ct} \geqslant 0 \quad \text{for all } c \in C \text{ and } t \in T. \tag{3''}$$

A matching is called *stable* if there exists no college c and student t who prefer to be matched to each other over the outcome they get from the given matching. We observe that the standard transformation of splitting each college c into q_c individual entities that are uniformly ordered by all students can be applied. For each such uniform order, one then gets a one-to-one correspondence of the stable matchings of the many-to-one (colleges to students) matching problem and the stable matchings of the corresponding one-to-one (men to women) matching problem, see Roth and Sotomayor (1990, Lemma 5.6). Thus, one can get a characterization of stable matchings via Theorem 1. We omit the details.

Acknowledgement

The author would like to thank Alvin E. Roth and two anonymous referees for many useful comments on an earlier version of this paper.

References

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

D. Gusfield and R.W. Irving, *The Stable Marriage Problem: Structure and Algorithms* (MIT Press, Cambridge, MA, 1989).

D.E. Knuth, *Mariages Stables* (Les Presses de l'Université Montréal, Montréal, 1976).

A.E. Roth, "The evolution of the labor market for medical interns and residents: a case study in game theory," *Journal of Political Economy* 92 (1984) 991-1016.

A.E. Roth, "The college admission problem is not equivalent to the marriage problem," *Journal of Economic Theory* 36 (1985) 277-288.

A.E. Roth and M.A.O. Sotomayor, *Two Sided Matchings - a Study in Game-Theoretic Modelling and Analysis* (Econometric Society Monograph Series, Cambridge University Press Cambridge, 1991).

J.H. Vande Vate, "Linear programming brings marital bliss," *Operations Research Letters* 8 (1989) 147-153.