Transformation from Arbitrary Matchings
to Stable Matchings

AKIHISA TAMURA

Department of Information Sciences, Tokyo Institute of Technology,
2-12-1 Oh-okayama, Meguro-ku, Tokyo 152, Japan

Received September 7, 1990

D. E. Knuth (1976, "Mariages stables," Presses Univ. Montréal, Montreal) conjectured that any matching can be transformed to some stable matching by a sequence of b-interchanges. Given a matching $M$ and a blocking pair $(m, w)$ for $M$, a b-interchange for $M$ by $(m, w)$ is defined as a transformation from $M$ to a matching obtained by replacing two pairs $(m, p_M(m))$ and $(p_M(w), w)$ in $M$ with $(m, w)$ and $(p_M(w), p_M(m))$. In this paper, we give a counter-example in which some matching cannot be transformed to any stable matching by b-interchanges. However, any matching can be transformed to some stable matching by using b-interchanges and identifying special cycling. We also give an algorithm to find either such cycling or a stable matching. © 1993 Academic Press, Inc.

1. INTRODUCTION

In an instance of the stable marriage problem of size $n$, each of $n$ men and $n$ women has a list of all members, called a preference list, of the opposite sex in the order of preference. Person $p$ prefers $q$ to $r$ if and only if $q$ precedes $r$ on $p$’s preference list, which we write as $q <_p r$. If either $q = r$ or $q <_p r$ then we write $q \preceq_p r$. A matching is a set of $n$ disjoint couples of men and women. If man $m$ and woman $w$ are coupled in a matching $M$, then $m$ and $w$ are called partners in $M$, which we write as either $m = p_M(w)$, $w = p_M(m)$, or $(m, w) \in M$ according to convenience. Man $m$ and woman $w$ are said to be a blocking pair for a matching $M$ if $w <_m p_M(m)$ and $m <_w p_M(w)$. If there is no blocking pair for $M$ then the matching $M$ is called stable. For a given stable marriage instance of size $n$, the divorce digraph is defined as follows. The node set of this digraph is the set of $n!$ matchings. The digraph has a directed edge from a matching $M$ to a matching $M'$ if and only if there is a blocking pair $(m, w)$ for $M$ such that $M'$ is obtained from $M$ by replacing $(m, p_M(m))$ and $(p_M(w), w)$ with

* Supported by Grant-in-Aids for Co-operative Research (63490010) of the Ministry of Education, Science and Culture.
TRANSFORMATION TO STABLE MATCHINGS

We call such a replacement arising from a blocking pair \((m, w)\) a \(b\)-interchange by \((m, w)\), and we denote \(M'\) by \(binter(M, m, w)\).

Gale and Shapley [1] proved that there exists at least one stable matching for any stable marriage instance. From the result, there is at least one sink in any divorce digraph. In a book of Knuth [3] it is conjectured that there is a path from each node to a sink in the divorce digraph, in other words, any matching can be transformed to some stable matching by a sequence of \(b\)-interchanges (see also Gusfield and Irving [2]). In Section 2, we provide a counter-example in which some matching cannot be transformed to any stable matching by \(b\)-interchanges. In fact if size \(n \geq 4\), one can always find such an instance. However, any matching can be transformed to some stable matching by using \(b\)-interchanges and identifying a special cycle in the divorce digraph. In Section 3, we give an algorithm to find either such a cycle or a stable matching.

2. COUNTER-EXAMPLE

Let \({m_0, ..., m_{n-1}}\) and \({w_0, ..., w_{n-1}}\) denote the sets of \(n\) men and \(n\) women. For each person \(p\), \(1(p), 2(p), ..., n(p)\) denote the first, second, ..., \(n\)th person on \(p\)'s preference list, respectively.

In order to deny Knuth's conjecture, we consider a stable marriage instance of size \(n \geq 4\) in which for each man \(m_i\) and each woman \(w_i\),

\[
\begin{align*}
1(m_i) &= w_i, & 2(m_i) &= w_{i-2}, \\
3(m_i) &= w_{i+1}, & 4(m_i) &= w_{i-1}, & 5(m_i), ... & \text{: arbitrary,} \\
1(w_i) &= m_{i+1}, & 2(w_i) &= m_{i-1}, \\
3(w_i) &= m_i, & 4(w_i) &= m_{i+2}, & 5(w_i), ... & \text{: arbitrary}
\end{align*}
\]

(see Fig. 1). In this section indices \(i-2, i-1, i+1, i+2,\) etc., are taken modulo \(n\). Let \(\mathscr{S}(n)\) denote such an instance of size \(n\). We will prove that \(\mathscr{S}(n)\) is a counter-example for Knuth's conjecture.

In the section we consider matchings \(M\) such that \(p_M(m_i) = k(m_i)\) for some \(k = 1, 2, 3, 4\), for each man \(m_i\). Let \(\text{shift}(M)\) denote the set of \(n\) couples in which \(m_{i+1}\)'s partner is \(w_{j+1}\) if \(p_M(m_i) = w_j\) for \(i = 0, 1, ..., n-1\). Obviously \(\text{shift}(M)\) is a matching, i.e., \((m_{i+1}, w_{j+1}) \in \text{shift}(M)\) if and only if \((m_i, w_j) \in M\). From the definition of \(\mathscr{S}(n)\), the matching \(\text{shift}(M)\) is obtained from \(M\) by cyclically shifting women among men on men's lists, i.e., \(\text{shift}(M)\) is equal to the matching in which \(m_{i+1}\)'s partner is \(k(m_{i+1})\) if \(p_M(m_i) = k(m_i)\) for \(i = 0, 1, ..., n-1\) (see Figs. 2 and 3). In the figures, the underlined person in each person's list is his or her partner in the
### FIG. 1. Stable marriage instance $J(n)$.

<table>
<thead>
<tr>
<th>Men's Preference Lists</th>
<th>Women's Preference Lists</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_0 )</td>
<td>( w_0 ) w(<em>{n-2}) ( w_1 ) ( w</em>{n-1} ) ( \ldots )</td>
</tr>
<tr>
<td>( m_1 )</td>
<td>( w_1 ) ( w_{n-1} ) ( w_2 ) ( w_0 ) ( \ldots )</td>
</tr>
<tr>
<td>[ \vdots ]</td>
<td>[ \vdots ]</td>
</tr>
<tr>
<td>( m_{n-3} )</td>
<td>( w_{n-3} ) ( w_{n-5} ) ( w_{n-2} ) ( w_{n-4} ) ( \ldots )</td>
</tr>
<tr>
<td>( m_{n-2} )</td>
<td>( w_{n-2} ) ( w_{n-4} ) ( w_{n-1} ) ( w_{n-3} ) ( \ldots )</td>
</tr>
<tr>
<td>( m_{n-1} )</td>
<td>( w_{n-1} ) ( w_{n-3} ) ( w_0 ) ( w_{n-2} ) ( \ldots )</td>
</tr>
</tbody>
</table>

### FIG. 2. $M = \{(m_0, w_0), \ldots, (m_{n-1}, w_{n-2})\}$.

<table>
<thead>
<tr>
<th>Men's Preference Lists</th>
<th>Women's Preference Lists</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_0 )</td>
<td>( w_0 ) w(<em>{n-2}) ( w_1 ) ( w</em>{n-1} ) ( \ldots )</td>
</tr>
<tr>
<td>( m_1 )</td>
<td>( w_1 ) ( w_{n-1} ) ( w_2 ) ( w_0 ) ( \ldots )</td>
</tr>
<tr>
<td>[ \vdots ]</td>
<td>[ \vdots ]</td>
</tr>
<tr>
<td>( m_{n-4} )</td>
<td>( w_{n-4} ) ( w_{n-6} ) ( w_{n-3} ) ( w_{n-5} ) ( \ldots )</td>
</tr>
<tr>
<td>( m_{n-3} )</td>
<td>( w_{n-3} ) ( w_{n-5} ) ( w_{n-2} ) ( w_{n-4} ) ( \ldots )</td>
</tr>
<tr>
<td>( m_{n-2} )</td>
<td>( w_{n-2} ) ( w_{n-4} ) ( w_{n-1} ) ( w_{n-3} ) ( \ldots )</td>
</tr>
<tr>
<td>( m_{n-1} )</td>
<td>( w_{n-1} ) ( w_{n-3} ) ( w_0 ) ( w_{n-2} ) ( \ldots )</td>
</tr>
</tbody>
</table>

### FIG. 3. $\text{shift}(M) = \{(m_0, w_{n-1}), \ldots, (m_{n-1}, w_0)\}$.
corresponding matchings. The matching \( \text{shift}(M) \) is also equivalent to the matching obtained from \( M \) by cyclically shifting men among women on women's lists, in which \( w_{i+1} \)'s partner is \( k(w_{i+1}) \) if \( p_M(w_i) = k(w_i) \) for \( i = 0, 1, ..., n - 1 \). One may infer the following lemmas from Figs. 2 and 3.

**Lemma 2.1.** In a stable marriage instance \( J(n) \), a pair \((m_i, w_j)\) is a blocking pair for a matching \( M \) if and only if \((m_{i+1}, w_{j+1})\) is a blocking pair for \( \text{shift}(M) \).

**Proof.** Let \( p_M(m_i) = w_h \) and let \( p_M(w_j) = m_k \). Then \( p_{\text{shift}(M)}(m_{i+1}) = w_{h+1} \) and \( p_{\text{shift}(M)}(w_{j+1}) = m_{k+1} \). The following equivalence relations hold:

\[
(m_i, w_j) \text{ is a blocking pair for } M \iff
\]

\[
w_j < m_i w_h = p_M(m_i) \text{ and } m_i < w_j m_k = p_M(w_j)
\]

\[
w_{j+1} < m_{i+1} w_{h+1} \text{ and } m_{i+1} < w_{j+1} m_{k+1}
\]

\[
(m_{i+1}, w_{j+1}) \text{ is a blocking pair for } \text{shift}(M).
\]

[from the definition of \( J(n) \)]

This completes the proof. \( \square \)

**Lemma 2.2.** In a stable marriage instance \( J(n) \), for a matching \( M \) and a blocking pair \((m_i, w_j)\) for \( M \),

\[
\text{shift}(\text{binter}(M, m_i, w_j)) = \text{binter}(\text{shift}(M), m_{i+1}, w_{j+1}).
\]

**Proof.** For a woman \( w \), we suppose that \( \text{shift}(w) \) denotes the woman \( w_{g+1} \) if \( w = w_g \). Let \( p_M(m_i) = w_h \) and let \( p_M(w_j) = m_k \). Since \((m_i, w_j), (m_k, w_h) \in \text{binter}(M, m_i, w_j), m_{i+1} \)'s partner in \( \text{shift}(\text{binter}(M, m_i, w_j)) \) is

\[
w_{j+1} \quad \text{if } l = i
\]

\[
w_{h+1} \quad \text{if } l = k
\]

\[
\text{shift}(p_M(m_i)) \quad \text{otherwise},
\]

for \( l = 0, ..., n - 1 \). On the other hand, from Lemma 2.1, \((m_{i+1}, w_{j+1})\) is a blocking pair for \( \text{shift}(M) \) and \((m_{i+1}, w_{h+1}), (m_{k+1}, w_{j+1}) \in \text{shift}(M)\).
Then $m_{i+1}$'s partner in $binter(shift(M), m_{i+1}, w_{j+1})$ is defined as above. Hence the equality holds. □

**Theorem 2.3.** For any size $n \geq 4$, in a stable marriage instance $J(n)$, there is a matching which cannot be transformed to any stable matching by $b$-interchanges.

**Proof.** We consider a matching

$$M_0 = \{(m_0, w_0), (m_1, w_1), \ldots, (m_{n-3}, w_{n-3}), (m_{n-2}, w_{n-1}), (m_{n-1}, w_{n-2})\}.$$  

From men's preferences, candidates of blocking pairs for $M_0$ are $(m_{n-2}, w_{n-2})$, $(m_{n-2}, w_{n-4})$, $(m_{n-1}, w_{n-1})$, $(m_{n-1}, w_{n-3})$ and $(m_{n-1}, w_0)$. See Fig. 2. The pair $(m_{n-1}, w_0)$, however, is the only blocking pair for $M_0$ since women $w_{n-2}$ and $w_{n-4}$ prefer their partners in $M_0$ to man $m_{n-2}$, women $w_{n-1}$ and $w_{n-3}$ prefer their partners in $M_0$ to man $m_{n-1}$, and $w_0$ prefers $m_{n-1}$ to $m_0$. Let $M'_0$ be the matching $binter(M_0, m_{n-1}, w_0)$, i.e.,

$$M'_0 = \{(m_0, w_{n-2}), (m_1, w_1), \ldots, (m_{n-3}, w_{n-3}), (m_{n-2}, w_{n-1}), (m_{n-1}, w_0)\}.$$  

From the preferences, there is also only one blocking pair $(m_{n-2}, w_{n-2})$ for $M'_0$. Let $M_1$ denote the matching $binter(M'_0, m_{n-2}, w_{n-2})$, i.e.,

$$M_1 = \{(m_0, w_{n-1}), (m_1, w_1), \ldots, (m_{n-3}, w_{n-3}), (m_{n-2}, w_{n-2}), (m_{n-1}, w_0)\}.$$  

The matching $M_1$ is uniquely determined from $M_0$ by two $b$-interchanges. On the other hand, $M'_1 = shift(M_0)$ (see Figs. 2 and 3). From Lemma 2.1, there is only one blocking pair $(m_0, w_1)$ for $M_1$. Lemma 2.2 implies that the matching $M'_1 = binter(M_1, m_0, w_1)$ identifies with $shift(M'_0)$. From the above facts, we can show that the sequence $M_0, M'_0, M_1, M'_1, \ldots$ of matchings is uniquely determined from $M_0$ by $b$-interchanges and $M_n = M_0$ holds. Therefore matchings $M_i$ and $M'_i$ for $i = 0, 1, \ldots, n-1$ cannot be transformed to any stable matching by $b$-interchanges. □

**Counter-example when $n = 4$.** The instance $J(4)$ is described in Fig. 4. This instance has 24 matchings and 5 stable matchings. We represent all matchings and blocking pairs for each matching in Fig. 5. The divorce digraph for $J(4)$ is drawn in Fig. 6. Matchings $M_1, M_8, M_{10}, M_{19}$ and $M_{24}$ are stable. There are paths from three matchings $M_6, M_{15}, M_{17}$ to the stable matching $M_1$. However, there is no path from any other unstable matchings to any one of the five stable matchings. For example, cycle $\{M_2, M_{16}, M_{22}, M_{12}, M_7, M_9, M_3, M_4\}$ of length 8 is obtained in the proof of Theorem 2.3 when size $n = 4$. 

TRANSFORMATION TO STABLE MATCHINGS

Men's Preference Lists

| m₀ | w₀ | w₂ | w₁ | w₃ |
| m₁ | w₁ | w₃ | w₀ | w₂ |
| m₂ | w₂ | w₀ | w₃ | w₁ |
| m₃ | w₃ | w₁ | w₀ | w₂ |

Women's Preference Lists

| w₀ | m₁ | m₃ | m₀ | m₂ |
| w₁ | m₀ | m₂ | m₁ | m₃ |
| w₂ | m₃ | m₁ | m₂ | m₀ |
| w₃ | m₀ | m₂ | m₃ | m₁ |

Fig. 4. Stable marriage instance \( \mathcal{S}(4) \).

<table>
<thead>
<tr>
<th>Matchings</th>
<th>Blocking Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>m₀ m₁ m₂ m₃</td>
<td>m₀ m₁ m₂ m₃</td>
</tr>
<tr>
<td>( M₁ ) (w₀ w₁ w₂ w₃)</td>
<td>( M₉ ) (m₃, w₀) ( \rightarrow M₁₆ )</td>
</tr>
<tr>
<td>( M₂ ) (w₀ w₁ w₃ w₂)</td>
<td>( M₃ ) (m₃, w₀) ( \rightarrow M₁₆ )</td>
</tr>
<tr>
<td>( M₃ ) (w₀ w₂ w₁ w₃)</td>
<td>( M₆ ) (m₂, w₀) ( \rightarrow M₈ )</td>
</tr>
<tr>
<td>( M₄ ) (w₀ w₂ w₃ w₁)</td>
<td>( M₈ ) (m₁, w₁) ( \rightarrow M₁ )</td>
</tr>
<tr>
<td>( M₅ ) (w₀ w₃ w₁ w₂)</td>
<td>( M₉ ) (m₁, w₂) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₆ ) (w₀ w₃ w₂ w₁)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₇ ) (w₁ w₀ w₂ w₃)</td>
<td>( M₉ ) (m₁, w₂) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₈ ) (w₁ w₀ w₃ w₂)</td>
<td>( M₉ ) (m₁, w₀) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₉ ) (w₁ w₂ w₀ w₃)</td>
<td>( M₉ ) (w₀, m₀) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₁₀ ) (w₁ w₂ w₃ w₀)</td>
<td>( M₉ ) (m₉, w₉) ( \rightarrow M₇ )</td>
</tr>
<tr>
<td>( M₁₁ ) (w₁ w₃ w₀ w₂)</td>
<td>( M₉ ) (m₉, w₉) ( \rightarrow M₇ )</td>
</tr>
<tr>
<td>( M₁₂ ) (w₁ w₃ w₂ w₀)</td>
<td>( M₉ ) (m₉, w₉) ( \rightarrow M₇ )</td>
</tr>
<tr>
<td>( M₁₃ ) (w₂ w₀ w₁ w₃)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₁₄ ) (w₂ w₃ w₁ w₀)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₁₅ ) (w₂ w₁ w₀ w₃)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₁₆ ) (w₂ w₁ w₃ w₀)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₁₇ ) (w₂ w₃ w₀ w₁)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₁₈ ) (w₂ w₃ w₁ w₀)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₁₉ ) (w₃ w₀ w₁ w₂)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₂₀ ) (w₃ w₀ w₂ w₁)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₂₁ ) (w₃ w₀ w₁ w₂)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₂₂ ) (w₃ w₁ w₀ w₂)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₂₃ ) (w₃ w₁ w₂ w₀)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
<tr>
<td>( M₂₄ ) (w₃ w₂ w₁ w₀)</td>
<td>( M₉ ) (m₁, w₁) ( \rightarrow M₉ )</td>
</tr>
</tbody>
</table>

Fig. 5. Matchings and blocking pairs in \( \mathcal{S}(4) \).
3. Algorithm for Finding a Stable Matching from a Matching

This section provides an algorithm which we call a \textit{b-interchange algorithm}, for finding either a directed path from a given matching $M_0$ to some stable matching or a cycle in the divorce digraph. A cycle is defined as a sequence of 3-tuples $(M_1, m_1, w_1), ..., (M_n, m_n, w_n)$ consisting of matchings $M_i$ and blocking pairs $(m_i, w_i)$ for $M_i$ such that $(M_1, m_1, w_1) = (M_n, m_n, w_n)$ and $M_{i+1} = \text{binter}(M_i, m_i, w_i)$ for $i = 1, ..., n - 1$. If a directed cycle $C$ is output by the algorithm, we can construct a matching $M$ from $C$ such that $bp(M)$ is a proper subset of $bp(M_0)$, where $bp(M)$ denotes the set of all blocking pairs for $M$. Hence, by iteratively using the algorithm, one can always arrive at some stable matching from an arbitrary matching.

For person $q$, let $bbp_M(q)$ denote the best person for $q$ among the set $\{p | (p, q) \text{ or } (q, p) \text{ is a blocking pair for } M\}$ if the set is nonempty; otherwise $bbp_M(q) = \text{nil}$. Then the $b$-interchange algorithm can be expressed as follows.
Algorithm 3.1 ($b$-interchange algorithm)

\[\text{(input, output): (a matching } M, \text{ some stable matching or cycle)}\]

\text{Step 0: } M_1 := M, i := 1;

\text{Step 1: if } M_i \text{ is stable then output } M_i \text{ and stop;}

\text{Step 2: while } \text{bbp}_M(M_i) \neq \text{nil} \text{ or } \text{bbp}_M(M_i) \neq \text{nil} \text{ do begin}

\text{Step 1: if } M_i \text{ is stable then output } M_i \text{ and stop;}

\begin{align*}
\text{Step 2: while } & \text{bbp}_M(M_i) \neq \text{nil} \text{ or } \text{bbp}_M(M_i) \neq \text{nil} \text{ do begin} \\
& p := \text{bbp}_M(M_i) \text{ such that } \text{bbp}_M(p) \neq \text{nil}; \\
& \text{if } p = \text{bbp}_M(M_i) \text{ then} \\
& \quad m_i := \text{bbp}_M(M_i); \quad w_i := \text{bbp}_M(M_i); \\
& \quad \text{marries happily in } M_{i+1} \\
& \quad \text{marries unhappily in } M_{i+1} \\
& \quad \text{else } \{ p = \text{bbp}_M(M_i) \} \\
& \quad \quad w_i := \text{bbp}_M(M_i); \\
& \quad \quad m_i := \text{bbp}_M(M_i); \\
& \quad \quad \text{marries happily in } M_{i+1} \\
& \quad \quad \text{marries unhappily in } M_{i+1} \\
& \quad \text{endif;}
\end{align*}

\text{if } (M_i, m_i, w_i) = (M_j, m_j, w_j) \text{ for some } j = 1, \ldots, i - 1 \text{ then}

\text{output the cycle } (M_j, m_j, w_j), \ldots, (M_i, m_i, w_i) \text{ and stop;}

\text{endif;}

\text{i := i + 1;}

\text{end \{while\};}

\text{goto Step 1;}

Given a matching } M \text{ and a blocking pair } (m, w) \text{ for } M, \text{ we say that pairs } (m, w) \text{ and } (p_M(m), p_M(w)) \text{ are happy and unhappy in the matching } binter(M, m, w), \text{ respectively, in the sense that unhappy persons were deserted by their partners and happy persons get better partners. It is also possible that an unhappy person gets a better partner. We call the pair defined in Step 1 unhappy for convenience. Informally, the b-interchange algorithm may be expressed as a sequence of b-interchanges determined by unhappy persons. At any point during the execution of Step 2, either an unhappy man } m_i \text{ or woman } w_i \text{ in a current matching } M_i \text{ determines the next blocking pair } (m_i, \text{bbp}_M(m_i)) \text{ or } (\text{bbp}_M(w_i), w_i). \text{ If } \text{bbp}_M(m_i) \neq \text{nil} \text{ and } \text{bbp}_M(w_i) \neq \text{nil}, \text{ then the algorithm has a flexible choice between } (m_i, \text{bbp}_M(m_i)) \text{ and } (\text{bbp}_M(w_i), w_i); \text{ however, there is no problem either way. We say that such blocking pairs determined by an unhappy man and woman are man-oriented and woman-oriented, respectively, if they exist. If such a blocking pair exists then a new matching } M_{i+1} \text{ is obtained by the b-interchange with the blocking pair. We call a b-interchange determined by a man-oriented (woman-oriented) blocking pair a man-oriented (woman-oriented) b-interchange. We note when the matching } M_{i+1} \text{ is obtained by a man-oriented b-interchange, the happy man is satisfied with the matching in the sense that there is no blocking pair containing him. Although the happy woman has a better partner, she may not be satisfied with the matching in the above sense. The same holds when } M_{i+1} \text{ is obtained by a}
woman-oriented b-interchange. The execution of Step 2 terminates when either a cycle is found or $bbp_{M_r}(\tilde{m}_i) = bbp_{M_r}(\tilde{w}_i) = \text{nil}$.

We first prove that Step 2 of the b-interchange algorithm reduces the number of blocking pairs if it find no cycle in the divorce digraph. Let $M_s$ be a matching just before the execution of the while statement in Step 2 and let $M_t$ be a matching just after the execution. For a matching $M_k$ ($k = s$, ..., $t$) at any point during the execution, the following lemma holds.

**Lemma 3.1.** Let $(m, w)$ be a pair satisfying at least one of the following conditions:

1. $m$ or $w$ is unhappy,
2. $(m, w)$ is not a blocking pair, i.e., $p_{M_k}(m) \preceq_m w$ or $p_{M_k}(w) \preceq_w m$ for $M_s$.

Then for any $k = s + 1$, ..., $t$, $(m, w)$ satisfies (1) or (2) for $M_k$.

**Proof.** We will prove the assertion by induction on $k$. Assume that condition (1) holds for $M_k$. If $(m, w)$ is an unhappy pair in $M_k$ then either $m$ or $w$ is clearly unhappy in $M_{k+1}$. Since men and women are symmetric, we consider the case when $m$ is unhappy and $w$ is not unhappy in $M_k$. If $M_{k+1}$ is obtained by the woman-oriented b-interchange then $m$ is also unhappy in $M_{k+1}$. So we suppose that $M_{k+1}$ is obtained by the man-oriented b-interchange determined by man $m$, below. If $(m, w)$ is a blocking pair for $M_k$, then $m$ has the partner $bbp_{M_k}(m)$ in $M_{k+1}$ with $bbp_{M_k}(m) \preceq_m w$, and hence, condition (2) holds for $M_{k+1}$. Suppose that $(m, w)$ is not a blocking pair for $M_k$, i.e., conditions (1) and (2) hold for $M_k$. If $p_{M_k}(m) \preceq_m w$ then $m$ has the partner $bbp_{M_k}(m)$ in $M_{k+1}$ with $bbp_{M_k}(m) \preceq_m p_{M_k}(m) \preceq_m w$; otherwise $p_{M_{k+1}}(w) = p_{M_k}(w) \preceq_w m$ since $M_{k+1}$ is obtained by the man-oriented b-interchange and since $(m, w)$ does not block $M_k$. Therefore condition (2) holds for $M_{k+1}$.

To complete the proof we consider the case when condition (2) holds for $M_k$ but (1) does not. Without loss of generality, we suppose that $p_{M_k}(m) \preceq_m w$. We can consider three possibilities: $m$ becomes unhappy in $M_{k+1}$, $m$ becomes happy in $M_{k+1}$ and the remaining case. In the first case, condition (1) holds for $M_{k+1}$. The second case implies that $p_{M_{k+1}}(m) \preceq_m p_{M_k}(m) \preceq_m w$, i.e., condition (2). In the third case, $p_{M_{k+1}}(m) = p_{M_k}(m) \preceq_m w$ holds.

**Lemma 3.2.** If no cycle is found in Step 2 then the set $bp(M_t)$ is a proper subset of $bp(M_s)$, i.e., $bp(M_s) \subset \not= bp(M_t)$.

**Proof.** From Lemma 3.1, if $(m, w)$ is not a blocking pair for $M_s$, then at least one of $m$ or $w$ is unhappy in $M_t$, or $(m, w)$ is not a blocking pair.
for $M_t$. If $m$ or $w$ is unhappy in $M_t$ then $(m, w)$ does not block $M_t$ since $bp_{M_t}(\tilde{m}_t) = bp_{M_t}(\tilde{w}_t) = nil$. Then $bp(M_t) \subseteq bp(M_s)$ holds.

Let $(m_s, w_s)$ be the initial blocking pair for $M_s$. Then either $m_s$ or $w_s$ is unhappy in $M_s$ (we recall that $(m_s, \tilde{w}_s)$ is an unhappy pair in $M_s$). In particular, from Lemma 3.1 and the fact that $bp_{M_t}(\tilde{m}_t) = bp_{M_t}(\tilde{w}_t) = nil$, as above, $(m_s, w_s)$ in $bp(M_s)$ does not block $M_t$. Hence $bp(M_t) \subseteq bp(M_s)$.

Lemma 3.2 says that the set of blocking pairs shrinks after Step 2 when no cycle is found.

We next consider the case when the b-interchange algorithm outputs a cycle $C$. In the rest of this section, we suppose that the cycle is defined as

$$C = \{(M_l, m_l, w_l), ..., (M_{n-1}, m_{n-1}, w_{n-1}), (M_n, m_n, w_n)\}.$$ 

We will show that one can obtain a matching $M$ such that $bp(M) \subseteq bp(M_s)$ by using the cycle $C$. Since men and women are symmetric, we will only prove assertions for men in the lemmas below.

Let $C_m$ and $C_w$ denote the sets of men and women whose partners are interchanged during the cycle $C$, respectively. Then the following lemma holds.

**Lemma 3.3.** If the b-interchange algorithm outputs a cycle $C$, then $C_m = \{m_l, ..., m_{n-1}\}$ and $C_w = \{w_l, ..., w_{n-1}\}$.

**Proof.** Obviously, $C_m \supseteq \{m_l, ..., m_{n-1}\}$. Assume on the contrary that there exists a man $m \in C_m \setminus \{m_l, ..., m_{n-1}\}$. Then, $m$ becomes unhappy at some point during $C$. After this point, there is no man-oriented b-interchange by $m$, because $m \notin \{m_l, ..., m_{n-1}\}$, which means that after this point the only unhappy man is $m$. So if any man ($\neq m$) changes his partner, he obtains a new partner whom he prefers to his old one. Because cycling means not only $M_i = M_n$ but also that the unhappy couple in the matching is the same, no cycle can occur under the assumption. Hence $C_m = \{m_l, ..., m_{n-1}\}$.

For each man $m \in C_m$, let $bbp_C(m)$ denote the best woman for $m$ among the set of women with whom $m$ causes b-interchanges during the cycle $C$, i.e., among the set $\{w_i | (m, w_i) = (m_l, w_l), i = l, ..., n-1\}$. From Lemma 3.3, $bbp_C(m)$ is well-defined. We define $bbp_C(w)$ for each woman $w \in C_w$ in the same way.

**Lemma 3.4.** If man $m \in C_m$ has a partner in some matching $M_i$ ($i = l, ..., n-1$) whom he prefers to $bbp_C(m)$ then $m$ is unhappy in $M_i$. The same also holds for women in $C_w$. 

Proof. Man $m$ changes his partner if and only if he becomes either happy or unhappy. From the definition of $bbp_c(m)$, if he becomes happy then his partner is $bbp_c(m)$ or below. Thus the assertion holds. 

**LEMMA 3.5.** For $m_1, m_2 \in C_m$ and $w_1, w_2 \in C_w$, 

$$m_1 \neq m_2 \Rightarrow bbp_c(m_1) \neq bbp_c(m_2),$$

$$w_1 \neq w_2 \Rightarrow bbp_c(w_1) \neq bbp_c(w_2).$$

**Proof.** Assume on the contrary that there exist two men $m_1, m_2 \in C_m$ with $bbp_c(m_1) = bbp_c(m_2) = w$. In addition, we suppose the following:

1. Let the cycle $C$ consist of matchings $M_1, ..., M_k$;
2. $m_1 < w m_2$ (w likes $m_1$ better than $m_2$);
3. $j = \min\{i \in [1, k-1] \mid M_i$ is obtained by a b-interchange with $(m_2, w)\}$;
4. $(m_1, w) \in M_i$ and $(m_1, w) \notin M_i$ for $i = 2, ..., j - 1$.

From assumption (4), $w$ is happy or unhappy in $M_2$. Since $bbp_c(m_1) = w$, $w$ must marry happily. Then $w$ has a partner in $M_2$ whom she prefers to $m_1$ and hence to $m_2$ also from assumption (2). Since $w$ and $m_2$ marry happily in $M_j$, she must be unhappy in some matching $M_i$ ($i = 3, ..., j - 1$). Let $h$ be the maximum in $\{3, ..., j - 1\}$ such that $w$ is unhappy in $M_h$. Since $(m_1, w) \notin M_h$, $m_1$ is not unhappy in $M_h$. From Lemma 3.4, $w < m_1 \rho M_h(m_1)$ holds. Then $w$ has a partner in $M_{h+1}$ whom she prefers to $m_1$ since $M_{h+1}$ is initiated by $w$ from the maximality of $h$. From the definitions of $h$ and $j$, $w$ prefers her partner in $M_j$, namely $m_2$, to $m_1$. This is a contradiction. 

For a cycle $C$ found by the b-interchange algorithm, let $M'^m_C$ be the set of man–woman pairs defined by

$$(m, bbp_c(m)) \in M'^m_C \quad \text{if} \quad m \in C_m$$

$$(m, \rho M(m)) \in M'^m_C \quad \text{if} \quad m \notin C_m \text{ for } M \in C.$$  

The set $M'^m_C$ of man–woman pairs can be defined similarly. Lemmas 3.3 and 3.5 guarantee that each of the sets $M'^m_C$ and $M'^w_C$ form a matching. Matchings $M'^m_C$ and $M'^w_C$ may not be obtained from the initial matching $M_s$ by a sequence of b-interchanges. By considering $M'^m_C$ or $M'^w_C$, we can reduce the set of blocking pairs. We will prove that $bp(M'^m_C), bp(M'^w_C) \subseteq bp(M_s)$, below.

**LEMMA 3.6.** For any $m \in C_m$, let $M_{i+1}$ be a matching obtained by a b-interchange determined by $(m, bbp_c(m))$. Then $w = bbp_c(m)$ is unhappy in
TRANSFORMATION TO STABLE MATCHINGS

$M_i$, so $(m, w)$ is a woman-oriented blocking pair. Similarly, for any $\bar{w} \in C_w$, $(bbp_C(\bar{w}), \bar{w})$ is a man-oriented blocking pair for $M_j$ if $M_{j+1}$ is obtained by a $b$-interchange with $(bbp_C(\bar{w}), \bar{w})$.

Proof. First we show that if $w$ is not unhappy in a matching then her partner in the matching is either $m$ or a man whom she prefers to $m$. Assume that $w$ is unhappy in a matching $M_M$ and marries happily in $M_{k+1}$. If $m$ is also unhappy in $M_k$, then $w$ has a partner in $M_{k+1}$ whom she prefers to $m$. Suppose that $m$ is not unhappy in the matching $M_k$. From Lemma 3.4, $m$ prefers $w = bbp_C(m)$ to $p_{M_k}(m)$. Then $w$'s partner in $M_{k+1}$ is either $m$ or a man whom she prefers to $m$. After $k + 1$, whenever $w$ is not unhappy, her partner is either $m$ or a man whom she prefers to $m$.

From the above fact, $w$ must be unhappy in $M_i$ because otherwise $(m, w)$ does not block $M_i$. 

We remark, from the proof of Lemma 3.6, that $m \in C_m$ is the worst partner for $bbp_C(m)$ among the set $\{m | (m_i, bbp_C(m)) \}$ is a man- or woman-oriented blocking pair during the cycle $C$ and that $w \in C_w$ is the worst for $bbp_C(w)$.

**Lemma 3.7.** Suppose that the $b$-interchange algorithm outputs a cycle $C$. Then, if $(m, w)$ is a blocking pair for $M_C^m$, then $m \notin C_m$ and $w \notin C_w$. The same is true for $M_C^w$.

Proof. Assume on the contrary that $w \in C_w$ holds. Let $\bar{m}$ be the man such that $bbp_C(\bar{m}) = w$. Since $(m, w)$ blocks $M_C^m$, $m \neq \bar{m}$ holds. We suppose that a matching $M_{i+1}$ is obtained from $M_i$ by a $b$-interchange determined by $(\bar{m}, w)$ during the cycle $C$. Then $\bar{m} <_w p_{M_i}(w)$. If $m$ was unhappy in $M_i$ then $w$ would prefer $\bar{m}$ to $m$, i.e., $(m, w)$ could not be a blocking pair for $M_C^m$, because $w$ is unhappy in $M_i$ from Lemma 3.6. Assume that $m$ is not unhappy in $M_i$. If $m \notin C_m$ then $p_{M_i}(m) = m$'s partner in $M_C^m$ and $w <_m p_{M_i}(m)$. If $m \in C_m$ then by Lemma 3.4 (which applies because $m \in C_m$), $p_{M_i}(m)$ is $bbp_C(m)$ or below, and hence, $w <_m p_{M_i}(m)$ since $(m, w)$ blocks $M_C^m$. Hence $(m, w)$ is a blocking pair for $M_i$ because $m <_w \bar{m} <_w p_{M_i}(w)$. But then $m$ should have been selected over $\bar{m}$ by $w$, a contradiction. Thus $w \notin C_w$ holds.

Assume on the contrary that $m \in C_m$ holds. From the above proof, $w \notin C_w$ holds. Let $M_{i+1}$ be a matching such that $m$ is happy in $M_{i+1}$ and unhappy in $M_i$. Then $bbp_C(m) \leq_m p_{M_{i+1}}(m) <_m p_{M_i}(m)$. Since $(m, w)$ is a blocking pair for $M_C^m$, $w <_m bbbp_C(m)$. On the other hand, $m <_w p_{M_i}(w) = p_{M_{i+1}}(w)$ because $w \notin C_w$. Then $(m, w)$ blocks $M_i$. But $m$ should have selected $w$ or a woman whom he prefers to $w$ as a partner in $M_{i+1}$. This is a contradiction.

Hence $m \notin C_m$ and $w \notin C_w$. 

Lemma 3.7 immediately implies that $bp(M'^*), bp(M^w) \subseteq bp(M)$ for any matching $M$ during the cycle $C$. One can prove a stronger result from Lemmas 3.1 and 3.7.

**LEMMA 3.8.** If the b-interchange outputs a cycle $C$, then

$$bp(M'^*), bp(M^w) \subseteq bp(M_s),$$

where $M_s$ is a matching just before the execution of Step 2.

**Proof.** Let $M$ be a matching during the cycle $C$ and let $(m, w)$ be a pair not blocking $M_s$. By Lemma 3.1, $m$ or $w$ is unhappy in $M$, or $(m, w)$ is not a blocking pair for $M$. In the first case, $m \in C_m$ or $w \in C_w$. From Lemma 3.7, $(m, w)$ is not a blocking pair for $M'^*$. If $m \notin C_m$, $w \notin C_w$, and $(m, w)$ does not block $M$ then $(m, w)$ is not a blocking pair for $M'^*$ since $p_M(m) = p_{M'^*}(m)$ and $p_M(w) = p_{M'^*}(w)$. Then $bp(M'^*) \subseteq bp(M_s)$.

Let $(m_s, w_s)$ be the initial blocking pair for $M_s$. That is, $m_s$ or $w_s$ is unhappy in $M_s$. We can show that $(m_s, w_s) \notin bp(M'^*)$ by using Lemmas 3.1 and 3.7, as above. Hence $bp(M'^*)$ is a proper subset of $bp(M_s)$. □

From Lemmas 3.2 and 3.8, one can obtain a matching $M$ with $bp(M) \subseteq bp(M_s)$ when the execution of Step 2 terminates. Therefore, by iteratively using the b-interchange algorithm, one can always arrive at some stable matching from an arbitrary matching.

**THEOREM 3.9.** Combining the b-interchange algorithm with transformations to $M'^*_{C_m}$ and $M'^*_{C_w}$, we can obtain some stable matching from an arbitrary matching.

**EXAMPLE.** We apply the b-interchange algorithm to the instance $J(4)$ in Fig. 4. Let $M_2 = \{(m_0, w_0), (m_1, w_1), (m_2, w_3), (m_3, w_2)\}$ be an input matching and let $(m_3, w_0)$ be the initial blocking pair. Then the algorithm outputs a cycle

$$C = \{(M_2, m_3, w_0), (M_{16}, m_2, w_2), (M_{22}, m_0, w_1), (M_{12}, m_3, w_3),
(M_7, m_1, w_2), (M_9, m_0, w_0), (M_3, m_2, w_3), (M_4, m_1, w_1), (M_2, m_3, w_0)\}.$$  

Any matching in this cycle cannot be transformed to any one of the five stable matchings (see Fig. 6). However, from the cycle $C$, one can construct two matchings

$$M'^*_{C_m} = \{(m_0, w_0), (m_1, w_1), (m_2, w_2), (m_3, w_3)\}$$

$$M'^*_{C_w} = \{(m_0, w_1), (m_1, w_2), (m_2, w_3), (m_3, w_0)\}.$$  

From Fig. 5, these are stable.
Remarks. It is not certain whether our algorithm for transforming a given matching to some stable matching terminates in polynomial time. From Lemma 3.3, the length \( k - 1 \) of a cycle \( C = (M_1, m_1, w_1), \ldots, (M_k, m_k, w_k) \) is at least \( 2 \times |C_m| = 2 \times |C_w| \). However, there is an indication that the length equals \( 2 \times |C_m| \) by our experiments for small stable marriage instances.

Acknowledgments

I thank Professor Donald E. Knuth and the referees for their valuable comments, which greatly improved the presentation of this paper.

References