<table>
<thead>
<tr>
<th>Title</th>
<th>On the NP-Completeness of weakly stable marriage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Miyazaki, Shuichi; Iwama, Kazuo</td>
</tr>
<tr>
<td>Citation</td>
<td>電子情報通信学会技術研究報告 信学技報 (1998), 98(432): 33-40</td>
</tr>
<tr>
<td>Issue Date</td>
<td>1998-11-20</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/227134">http://hdl.handle.net/2433/227134</a></td>
</tr>
<tr>
<td>Rights</td>
<td>© 1998 by the Institute of Electronics, Information and Communications Engineers (IEICE)</td>
</tr>
<tr>
<td>Type</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>

Kyoto University
条件を緩和した安定結婚問題のNP完全性について

宮崎 修一 岩間 一雄
京都大学 大学院情報学研究科
〒 606-8501 京都市左京区吉田本町
Tel 075-753-5392
Fax 075-753-5379
{shuichi,iwama}@uis.kyoto-u.ac.jp

あらまし
安定結婚問題は、入力としてN人ずつの男女と各個人の異性に対する希望リスト（異性N人の順序付け）が与えられ、安定な組のペアを探す（安定なN組のペアが存在するかどうかを問う）問題である。この問題の自然な拡張として、(i) 希望リストに異性N人の全員を書かなくても良いこと、(ii) 希望の順序付けに不順序を許すものがあるが、いずれも解の存在は多項式時間で判定できることが知られている。本研究では、(i) および(ii)を同時に許した場合に、本問題がNP完全になることを示す。また、本問題に関連した問題のNP完全性も示す。

キーワード 安定結婚問題、順序付け、学生配置問題、NP完全性

On the NP-Completeness of Weakly Stable Marriage

Shuichi Miyazaki Kazuo Iwama
Graduate School of Informatics, Kyoto University
Yoshida-Honnachi, Sakyō-ku, Kyoto 606-8501
Phone +81-75-753-5392
Fax +81-75-753-5379
{shuichi,iwama}@uis.kyoto-u.ac.jp

Abstract
The original stable marriage problem requires both each man and woman to submit a complete and strict order of his/her preference, i.e., a total order list of, say, 100 people, if there are 100 men and women. This is obviously unrealistic often in practice, and several relaxations have been proposed, including the following two common ones: One is to allow an incomplete list, i.e., a man is allowed to accept only a subset of the women and vice versa. The other is to allow a "weaker" preference list such as a list including ties and partial orders. A little surprisingly, it is known that both problems can still be solved in polynomial time. In this paper, we show that the problem becomes NP-hard, if we allow both relaxations (incomplete lists and partial orders) at the same time.

key words stable marriage problem, partial order, student assignment problem, NP-completeness
1 Introduction

An instance of the stable marriage problem [GI89] consists of $N$ men and $N$ women. Each person has his/her strictly ordered preference list containing all the members of the opposite sex. A matching $M$ is one-one correspondence between the men and the women. If man $m$ and woman $w$ are matched in $M$, we say $m$ and $w$ are partners in $M$, and we write $p_M(m) = w$ and $p_M(w) = m$. A man $m$ and a woman $w$ is called a blocking pair for matching $M$, if $m$ and $w$ are not partners in $M$, but $m$ prefers $w$ to $p_M(m)$ and $w$ prefers $m$ to $p_M(w)$. If there is no blocking pair for $M$, then $M$ is called stable. If a matching $M$ contains all the $2N$ members, it is called complete. Stable marriage problem was first studied by Gale and Shapley [GS62]. They showed that there always exists a complete stable matching in any instance and showed $O(n^2)$ time algorithm (so-called Gale-Shapley algorithm) to find one.

However, considering practical applications, the restriction for preference lists appears to be too strict in many occasions because of the total order and complete list of preferences. In this context, there are some extensions of this problem. The popular ones are (i) the stable marriage problem with unacceptable partners [GI89], and (ii) the stable marriage problem with indifference [GI89, Ir94]. In the first extension, each person is allowed to declare one or more unacceptable partners. Thus each person’s preference list may be incomplete. Gale and Sotomayor [GS85] studied this extension and showed that a maximum stable matching is found in polynomial time using Gale-Shapley algorithm slightly modified to treat the unacceptable case (see also [GI89]). Thus the problem does not become essentially harder.

In the second extension, each person is allowed to have a preference list with ties, or more generally, a preference list of partial order. $m$ and $w$ are called a blocking pair if $m$ and $w$ are not partners in $M$, but $m$ strictly prefers $w$ to $p_M(m)$ and $w$ strictly prefers $m$ to $p_M(w)$ in the partial order list. A matching containing no such blocking pair is called weakly stable. However, this extension also does not make the problem significantly difficult; it is known that there is a polynomial-time algorithm which determines whether there exists a complete stable matching (and finds one if exists) [GI89, Ir94]. In [Ir94], two other notions of stabilities are also introduced but the situation is still the same, i.e., the problem is solvable in polynomial time.

In this paper, we consider how the situation changes if we apply both extensions, i.e., incomplete preference lists and partial orders, at the same time. Our main purpose of this paper is to show that the situation does change, i.e., the problem is now NP-complete. Also, it is shown that the problem is also intractable for complete lists if the question is not the existence of complete matching but obtaining optimal (complete) matchings. The general perception on the stable marriage problem has been that it is basically not hard [GS82, GI89, Ir94]. (Two exceptions are mentioned in [Ro90], either of which is not the bipartite setting.) Our results in this paper could give some change to this common perception.

In this paper, we introduce the student assignment problem, which is a special case of the marriage problem. There are $N$ graduate students and $N$ professors, where each student has a preference list of professors and students are ordered according to the scores of the entrance examination. We are asked to assign each student to each professor under some stable condition. In Sec 2, we give formal definitions of several extensions of the stable marriage problem and student assignment problem, and show the relationship between them. Sec.3 is devoted for the proof of several intractability results of the marriage problem with unacceptable partners and partial order preference lists. The main proof is given against the student assignment problem.

2 Stable Marriage and Student Assignment Problems

Recall that the original stable matching problem requires each person's preference list to be complete (i.e., all the members of the opposite sex must be included) and to be a total order. We focus on the three possible relaxations concerning these two restrictions. All these problems [including the
original one) ask if there exists a complete stable matching. In each problem, we are given \( N \) men, \( N \) women and each person's preference list. The condition for a pair of a man \( m \) and a woman \( w \) to be a blocking pair is the same for all: \( m \) and \( w \) is a blocking pair for matching \( M \), if \( m \) and \( w \) are not partners in \( M \), but \( m \) strictly prefers \( w \) to \( p_M(m) \) and \( w \) strictly prefers \( m \) to \( p_M(w) \). Thus to summarize these problems, it suffices to specify conditions for preference lists:

**Problem:** Stable Marriage Problem with Complete List and Total Order (SMP-CLTO).
Preference List: Each person's preference list must be complete and must be a total order.

**Problem:** Stable Marriage Problem with Incomplete List and Total Order (SMP-ILTO).
Preference List: Each person's preference list may be incomplete but must be a total order.

**Problem:** Stable Marriage Problem with Complete List and Partial Order (SMP-CLPO).
Preference List: Each person's preference list must be complete but may be a partial order.

**Problem:** Stable Marriage Problem with Incomplete List and Partial Order (SMP-ILPO).
Preference List: Each person's preference list may be incomplete and may be a partial order.

The first three problems are known to be in P [GI89, IR94]. We later show the NP-completeness of the fourth one. If we consider a cost of a matching, we can consider another extension. Let \( m \) be a man and \( L = (w_1, w_2, \ldots, w_N) \) be his preference list. A **weighted** preference list is a combination of \( L \) and a weight function \( \sigma \) which maps each \( w_i \) into a real number such that if \( w_i \) precedes \( w_j \) in \( m \)'s list, then \( \sigma_m(w_i) \leq \sigma_m(w_j) \). Namely, the cost function must not contradict to the total order.

For \( m \) is matched to \( w_i \) in a matching \( M \), then \( m \)'s cost under \( M \) is defined to be \( \sigma_m(w_i) \). A similar cost function is also introduced for the women's preference lists. A cost of the matching \( M \) is the sum of the costs of all \( 2N \) persons.

**Problem:** \( k \)-SMP-CLTO.
Instance: \( N \) men, \( N \) women, a positive integer \( k \), each person's preference list which must be complete and must be a total order, and a cost function \( \sigma \).
Question: Is there a complete stable matching whose cost is equal to or smaller than \( k \)?

**Problem:** \( k \)-SMP-CLPO.
Instance: \( N \) men, \( N \) women, a positive integer \( k \), each person's preference list which must be complete but may be a partial order, and a cost function \( \sigma \) which does not contradict to the partial order.
Question: Is there a complete stable matching whose cost is equal to or smaller than \( k \)?

The first optimization problem is known to be in P [GI89], i.e., a stable matching with minimum cost can be obtained in polynomial time. We later show the NP-completeness of the second problem.

Finally let us consider the **Student Assignment Problem (SAP for short)** introduced in Sec.1: An instance of SAP consists of the same number \( N \) of students and professors, the **rank list** of students, and a **preference list** of each student. The rank list is a total order of students. If student \( s_i \) is in the \( j \)th position on the rank list, we write \( r(s_i) = j \). Each student has a preference list of a total order of (not necessarily all) professors. If professor \( a_j \) is in the \( i \)th position on \( s_i \)'s list, we write \( \delta_{s_i}(a_j) = i \). An **assignment** \( A \) is a bijection from students to professors; if student \( s_i \) is assigned to professor \( a_j \), we write \( A(s_i) = a_j \). Each \( s_i \) must be mapped to a professor who exists on \( s_i \)'s preference list. Suppose that, under an assignment \( A \), students \( s_i \) and \( s_j \) are assigned to professors \( a_i \) and \( a_m \), respectively, or \( A(s_i) = a_i \) and \( A(s_j) = a_m \). Then a pair of a student \( s_i \) and a professor \( a_m \) is called a **blocking pair** for \( A \) if (1) \( r(s_i) < r(s_j) \), (2) \( \delta_{s_i}(a_m) < \delta_{s_i}(a_i) \) and (3) \( \delta_{s_i}(a_m) \leq \delta_{s_i}(a_m) \). If there is no blocking pair for \( A \), then \( A \) is called **stable**. Here we consider only one version, i.e., the preference lists of students may be incomplete but must be a total order.

**Problem:** Student Assignment Problem (SAP).
Instance: \(N\) students, \(N\) professors, and each student's preference list which may be incomplete but must be a total order.

Question: Is there a stable assignment?

Lemma 1. SAP is a special case of SMP-ILPO.

Proof. We consider each professor's preference list as a partial order of students. First, professor \(a_i\)'s preference list does not include student \(s_j\) if \(a_i\) does not appear on \(s_j\)'s preference list. Then consider two students \(s_i\) and \(s_j\) included in \(a_i\)'s list. \(a_i\) strictly prefers \(s_i\) to \(s_j\) if and only if (i) \(r(s_i) < r(s_j)\), and (ii) \(\delta(s_i(a_i)) \leq \delta(s_j(a_i))\). (Blocking conditions (1) and (3) are combined into one, and reflected in a preference lists of professors.) It is not hard to see that the preference of each professor is a partial order of students, namely. A blocking pair then is a pair of a student and a professor, who are not currently matched but strictly prefer each other to their current partners.

3 Intractability Results

We first prove that SAP is NP-complete. It immediately implies that SMP-ILPO is NP-complete by Lemma 1. Also it is straightforward to imply that \(k\)-SMP-CLPO is NP-hard using some artificial weight function. We also discuss more natural weight functions.

Theorem 1. SAP is NP-complete.

Proof. It is easy to see that SAP is in NP: Given an assignment \(A\), one can check whether \(A\) is a bijection or not. Then check, for each pair of a student and a professor, whether it is a blocking pair. This can be done in polynomial time.

To show the NP-hardness, let us consider the following problem:

Problem: ONE-IN-THREE 3SAT.

Instance: 3CNF formula.

Question: Is there any truth assignment such that exactly one literal in each clause is true?

It is known that ONE-IN-THREE 3SAT remains NP-complete even if a 3CNF formula does not include negative literals [GJ79, Sc78]. So, we translate this restricted problem into SAP. Given \(f\), which is an instance of ONE-IN-THREE 3SAT, we construct an instance \(T(f)\) of SAP, namely, (1) the same number of students and professors, (2) a rank list of students, and (3) each student's preference list. Let \(n\) and \(m\) be the numbers of variables and clauses of \(f = C_1 \cdot C_2 \cdots C_m\), respectively. Let \(t_i\) be the number of appearances of \(x_i\) and \(t\) be the maximum number among \(t_1, t_2, \ldots, t_n\). \(T(f)\) consists of \(9m + 3n + t + 6\) students and the same number of professors. We first introduce the following \(9m + 3n + t + 6\) students:

Group (i): \(t + 6\) students \(s_{i1}, \ldots, s_{i6}\).

Group (ii): \(n\) students \(s_{21}, \ldots, s_{2n}\).

Group (iii): \(2n\) students \(s_{31}, \ldots, s_{3n}, s_{51}, s_{52}, s_{53}\).

Group (iv): \(6m\) students \(s_{61,i}, s_{62,i}, s_{63,i}\). For \(1 \leq i \leq n\) and \(1 \leq j \leq m\), students \(s_{61,i}^{+}\) and \(s_{62,i}^{-}\) exist if and only if clause \(C_j\) contains literal \(x_i\). (Recall that \(\overline{x_i}\) never appears in \(f\).) Since there are \(3m\) literals, \(6m\) students will be introduced.

Group (v): \(3m\) students \(s_{63,ij}\) for any \(1 \leq i \leq m\) and \(1 \leq j \leq 3\). Students \(s_{63,1i}, s_{63,2i}, s_{63,3i}\) correspond to clause \(C_i\).

Students of Group (i) are necessary for technical reason. A student \(s_{63}\) of Group (ii) and two students \(s_{61}^{+}\) and \(s_{62}^{-}\) of Group (iii) are associated with variable \(x_i\). Students \(s_{62,ij}^{+}\) and \(s_{62,ij}^{-}\) of Group (iv) are associated with literal \(x_i\) in clause \(C_j\). The same number \((9m + 3n + t + 6)\) of professors are divided into the following three categories:

1. \(t + 6\) professors \(g_{11}, \ldots, g_{16}\).
2. \(3n\) professors \(i\) and \(P_i\) \((1 \leq i \leq n)\). Three professors \(i\) and \(P_i\) are associated with variable \(x_i\).
Students of Group (ii) selects two students have the same value of i, then the student with smaller j is higher. As for Group -(i) students, each student selects only one professor, i.e (the 5th and 4th positions are always used without depending on respectively. (The 5th and 4th positions are filled with 9 lists using students 4 professors on the 1st through 8. Group (iv) Groups (i), (ii) and (iii) are determined as explicitly described above, e.g., r(s1) < r(s2). Within Group (iv), a student with smaller index i is higher in the rank. If two students have the same value of i, then the student having smaller j is higher. If both i and j are the same, the student with + is higher than the student with -. Within Group (v), a student with smaller i is higher. If two students have the same value of i, then the student with smaller j is higher. 

Secondly, we determine the rank of students as follows: Among the groups, the order goes as Group (i) highest, Group (ii) next, and so on, until Group (v) that is lowest. The order within Groups (i), (ii) and (iii) are determined as explicitly described above, e.g., r(s1) < r(s2). Within Group (iv), a student with smaller index i is higher in the rank. If two students have the same value of i, then the student having smaller j is higher. If both i and j are the same, the student with + is higher than the student with -. Within Group (v), a student with smaller i is higher. If two students have the same value of i, then the student with smaller j is higher. 

Finally, we construct each student’s preference list. For better exposition, we use an example of f, i.e., f = (x1 + x2 + x3)(x1 + x5 + x4), for which the preference list is turned out to be as illustrated in Table 2. As for Group -(i) students, each student selects only one professor, i.e., s1,i selects g1,i. Students of Group (ii) selects t + 8 professors in 1st to (t + 8) th positions of the preference lists. Professors on the 1st through (t + 6) th positions are g1,1 through g1,8, arbitrarily. s2,j selects professor i in the (t + 7) th position and selects professor i in the (t + 8) th position, namely, δst,i(t) = t + 7 and δst,j(t) = t + 8.

Then we construct preference lists of Group-(iii) students. We show how to construct preference lists using students s3,1,1 and s3,2,1 for f0. (Again, see Table 2.) Recall that these two students are associated with variable x1. Students s3,1,1 and s3,2,1 write professor P1 in the 5th and the 4th positions, respectively. (The 5th and 4th positions are always used without depending on f.) Then, s3,1,1 writes professor i in the (t + 6)(= 8) th position and s3,2,1 writes T in the same position. Since x1 appears in clauses C1 and C2, four professors (1, 1), (1, 1), (1, 2) and (1, 2) have been introduced. s3,1,1 writes (1, 1) and (1, 2) in the 6th and 7th positions, respectively, and s3,2,1 writes (1, 1) and (1, 2) also in the 6th and 7th positions, respectively. Other positions are filled with g1,1 through g1,8 arbitrarily.

Now we move to Group (iv) students: We describe how to construct preference lists of s4,1,1 and s4,2,1 of Group (iv) in the example, who are associated with variable x1. Student s4,1,1 writes professor Q1,1 in the 5th position and writes (1, 1) in the same position as s3,1,1 wrote it. s4,2,1 writes Q1,1 in the 4th position and writes (1, 1) in the same position as s3,2,1 wrote it. Blanks are filled with professors g1,1 through g1,8 arbitrarily.

Recall that three students s5,1,1, s5,2,2 and s5,3,3 of Group (v) are associated with clause C1. We show how to construct preference lists of students s5,2,1, s5,2,2 and s5,2,3 associated with C1 = (x1 + x3 + x4) of f0. Since literals x1, x3 and x4 appear in C1, six professors (1, 1), (1, 2), (3, 2), (4, 2) and (4, 2) have been introduced. s5,2,1 writes (1, 2), (3, 2) and (4, 2) in this order from the 1st to the 3rd positions. Both s5,2,2 and s5,2,3 write (1, 2), (3, 2) and (4, 2) in this order.

The translation is completed. Table 2 shows the whole lists of T(f) of f. Next, we show a series of lemmas which summarize several conditions for an assignment A for T(f) to be a solution of the problem. We omit the proofs but the correctness can be easily seen. Recall that we denote A(s1) = a1 if student s1 is assigned to professor a1 under the assignment A.

**Lemma 2.** If an assignment A for T(f) is a solution, then A(s1) = g1,i (1 ≤ i ≤ t + 6). Namely, each student of Group (i) is assigned to the professor of his/her first position on the list.

**Lemma 3.** If an assignment A for T(f) is a solution, then for each i (1 ≤ i ≤ m), A(s2,i) = i or A(s2,i) = 7.

**Lemma 4.** Suppose that an assignment A for T(f) is a solution. Then for all 1 ≤ i ≤ n, (a) if A(s2,i) = i then A(s2,i) = P1 and A(s2,i) = 7. (b) Otherwise, i.e., if A(s2,i) = 7, then A(s2,i) = i and A(s2,i) = P1.

**Lemma 5.** Suppose that an assignment A for T(f) is a solution. Then following (a) and (b) hold for all i (1 ≤ i ≤ n). (a) If A(s2,i) = i then A(s2,i) = (i, j) and A(s2,i) = Q4,i for all j. (b) If A(s2,i) = 7 then A(s2,i) = Q4,i and A(s2,i) = (i, j) for all j.

**Lemma 6.** Suppose that an assignment A for T(f) is a solution. Then, for each i and j, one
of (i, j) and (j, j) is assigned to a student of Group (iv), and the other is assigned to a student of Group (v).

Let $C_i = \{x_{i1} + x_{i2} + x_{i3}\}$ be a clause in $f$, where $j_1 \leq j_2 \leq j_3$. Then recall that there are three students $s_{5,1}$, $s_{5,2}$ and $s_{5,3}$, and six professors $(j_1, j), (j_1, j), (j_2, j), (j_3, j), (j_3, j)$ and $(j_3, j)$ (see Table 1).

<table>
<thead>
<tr>
<th>$s_{5,1}$</th>
<th>$(j_1, j)$</th>
<th>$(j_2, j)$</th>
<th>$(j_3, j)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{5,2}$</td>
<td>$(j_1, j)$</td>
<td>$(j_2, j)$</td>
<td>$(j_3, j)$</td>
</tr>
<tr>
<td>$s_{5,3}$</td>
<td>$(j_1, j)$</td>
<td>$(j_2, j)$</td>
<td>$(j_3, j)$</td>
</tr>
</tbody>
</table>

Table 1: Preference lists of students associated to $C_i$

**Lemma 7.** Suppose that an assignment $A$ for $T(f)$ is a solution. Then, for all $j (1 \leq j \leq m)$, $s_{5,1}$, $s_{5,2}$ and $s_{5,3}$ are assigned in one of the following ways (a) and (b) and (c): (a) $A(s_{5,1}) = (j_1, j)$, $A(s_{5,2}) = (j_2, j)$ and $A(s_{5,3}) = (j_2, j)$, (b) $A(s_{5,1}) = (j_1, j)$, $A(s_{5,2}) = (j_1, j)$ and $A(s_{5,3}) = (j_3, j)$, (c) $A(s_{5,1}) = (j_1, j)$, $A(s_{5,2}) = (j_1, j)$ and $A(s_{5,3}) = (j_2, j)$.

**Lemma 8.** Suppose that an assignment $A$ for $T(f)$ is a solution. Then, for each $i$ and $j$, the following statements (a), (b) and (c) are true: (a) If $A(s_{5,1}) = (i, j)$, then $A(s_{5,1}) = (i, j)$. (b) If $A(s_{5,2}) = (i, j)$ then $A(s_{5,2}) = (i, j)$. (c) If $A(s_{5,3}) = (i, j)$ then $A(s_{5,3}) = (i, j)$.

Now we are ready to show the correctness of the reduction. To make the argument clear, we denote the literal $x_i$ in clause $C_j$ by $x_i^j$. Let us consider the following association rule between assignment for variables (and literals) of $f$ and assignment $A$ for students of $T(f)$: (1) Assign 1 to variable $x_i$ if and only if student $s_{5,1}$ is assigned to professor $i$, and assign 0 to variable $x_i$ if and only if student $s_{5,2}$ is assigned to professor $i$. (2) Assign 1 to literal $x_i^j$ if and only if student $s_{5,3}$ is assigned to professor $i$, and assign 0 to literal $x_i^j$ if and only if student $s_{5,3}$ is assigned to professor $i$. We will show the consistency, namely, under the assumption that $A$ is a solution for $T(f)$, if $x_i = 1$, then $x_i^j = 1$ for all $j$, and if $x_i = 0$, then $x_i^j = 0$ for all $j$.

Now we are ready to show the correctness of the reduction. To make the argument clear, we denote the literal $x_i$ in clause $C_j$ by $x_i^j$. Let us consider the following association rule between assignment for variables (and literals) of $f$ and assignment $A$ for students of $T(f)$: (1) Assign 1 to variable $x_i$ if and only if student $s_{5,1}$ is assigned to professor $i$, and assign 0 to variable $x_i$ if and only if student $s_{5,2}$ is assigned to professor $i$. (2) Assign 1 to literal $x_i^j$ if and only if student $s_{5,3}$ is assigned to professor $i$, and assign 0 to literal $x_i^j$ if and only if student $s_{5,3}$ is assigned to professor $i$. We will show the consistency, namely, under the assumption that $A$ is a solution for $T(f)$, if $x_i = 1$, then $x_i^j = 1$ for all $j$, and if $x_i = 0$, then $x_i^j = 0$ for all $j$.

Suppose $x_i = 1$. Then $A(s_{5,1}) = i$ by the association rule. Suppose, for some $j$ such that literal $x_i$ exists in clause $C_j$, $s_{5,2}$ or $s_{5,3}$ is assigned to professor $(i, j)$. Then, by (b) or (c) of Lemma 8, $A(s_{5,3}) = i$, which is a contradiction, since $s_{5,3}$ is assigned to both $i$ and $i$ under $A$. Thus if $x_i = 1$ then $x_i^j = 1$ for all $j$. The other one, i.e., for $x_i = 0$, can be shown similarly.

Suppose there exists a solution $A^*$ for $T(f)$. By Lemma 3, either $A^*(s_{5,1}) = i$ or $A^*(s_{5,2}) = i$ for all $i$. Then we determine the assignment to variables of $f$ using the association rule. If $A^*(s_{5,2}) = i$ then $x_i = 0$, otherwise, i.e., if $A^*(s_{5,2}) = i$, then $x_i = 1$. We show that this is a solution for $f$. Let, for $1 \leq j \leq m$, the $j$th clause of $f$ be $C_j = (x_{j1} + x_{j2} + x_{j3})$. Then preference lists of three students associated to $C_j$ are the ones described in Table 1. By Lemma 7, these students must be assigned in either (a) or (b) or (c) in Lemma 7. Suppose it is (a), namely, $A^*(s_{5,1}) = (j_1, j)$, $A^*(s_{5,2}) = (j_2, j)$ and $A^*(s_{5,3}) = (j_2, j)$.

Conversely, suppose there exists a solution for $f$ of ONE-IN-THREE 3SAT. Then, again using the association rule, we determine assignment for students of Group (ii) and (v). Students of Groups (i), (iii) and (iv) are automatically assigned. The fact that this assignment is a solution for $T(f)$ can be easily seen from the lemma we have already shown.

**Corollary 1.** SMP-ILPO is NP-complete.

**Proof.** Membership in NP is obvious. We have shown, in Lemma 1, that SMP-ILPO is a special case of SAP.

---
Corollary 2. k-SMP-CLPO is NP-complete.

Proof. Membership in NP is obvious. We will give a reduction from SMP-ILPO. We first make each man m’s preference list complete: We add a woman w (who did not exist in m’s original list) to m’s list in such a way that m does not prefer w to any other women who did originally exist in the list. We define a weight function in the following way: The weight of each woman who was in m’s original list is 0, and the weight of each woman newly added is 1. We do the similar process to each woman’s list also. Then it is easily seen that there exists a complete stable matching for SMP-ILPO if and only if there exists a stable matching for k-SMP-CLPO whose cost is 0. □

Let us consider a much more natural cost function, denoted by C, than the one used in Corollary 2. Define si’s cost under A be $C_A(s_i) = \delta_i(A(s_i))$, namely, the position of the professor whom $s_i$ is assigned to. (δ is defined in Sec.2) The cost of an assignment A is the sum of the costs of all the students under A. As one can see easily, this cost function can be used for SMP-ILPO if the preference lists of either men or women are total order.

Problem: k-SAP.

Instance: N students, N professors, a positive integer $k$, and each student’s preference list which must be complete and must be a total order.

Question: Is there a complete stable assignment whose cost is equal to or smaller than $k$?

Theorem 2. k-SAP is NP-complete.

Proof (Sketch). Membership in NP is obvious. We reduce $T(f)$ obtained in the proof of Theorem 1 into an instance $T'(T(f))$ of k-SAP. We introduce a large number, say $n^2m^2$, of Group-(i) students $s_{i,j}$ and professors $g_{i,j}$ and make each student’s preference lists complete by first adding $g_{i,j}$ to the tail of the list and then other remaining professors. It can be seen that there is no complete matching in $T(f)$ if and only if at least one student in $T'(T(f))$ is assigned to a professor of very low position on the list. If we set appropriate $k$, only the cost of this student exceeds $k$. □

References


<table>
<thead>
<tr>
<th>Table 2: Reference list of students of T(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Katsuyuki Nakajima</td>
</tr>
<tr>
<td>Keiichi Iida</td>
</tr>
<tr>
<td>Yoichiro Tanaka</td>
</tr>
</tbody>
</table>

**The Complexity of Schema Matching**

Abstract

In this paper, we study the schema matching problem, which is the special case of the second-order matching problem in the framework of formalism. A schema matching problem is defined as a problem of finding a mapping between two sets of data. The mapping is obtained by solving a constrained optimization problem. The optimization problem is formulated as a quadratic programming problem. The problem is NP-complete. We then present a polynomial-time algorithm for solving the problem. The algorithm is based on a greedy approach. The algorithm works as follows:

1. Initialize the mapping by setting each element of one set to be matched with an element of the other set.
2. For each unmatched pair of elements, compute the cost of matching them. The cost is defined as the sum of the costs of matching each element of the pair.
3. Select the pair with the smallest cost and match them.
4. Repeat steps 2 and 3 until all elements are matched.

The algorithm returns a mapping that minimizes the total cost of the mapping. The algorithm runs in polynomial time.

Furthermore, we discuss the schema matching problem including the under-constrained matching problem, which is an extension of the problem.

Reference