Egerváry Research Group
on Combinatorial Optimization

Technical reportS

TR-2017-04. Published by the Egerváry Research Group, Pázmány P. sétány 1/C, H–1117, Budapest, Hungary. Web site: www.cs.elte.hu/egres . ISSN 1587–4451.

Finding strongly popular b-matchings in bipartite graphs

Tamás Király and Zsuzsa Mészáros-Karkus

March 2017

Finding strongly popular b-matchings in bipartite graphs

Tamás Király[⋆] and Zsuzsa Mészáros-Karkus[⋆⋆]

Abstract

The computational complexity of the bipartite popular matching problem depends on the type of indifference allowed in the preference lists. If one side has strict preferences while nodes on the other side are indifferent (but prefer to be matched), then a popular matching can be found in polynomial time [Cseh, Huang, Kavitha, 2015]. However, as the same paper points out, the problem becomes NP-complete if one side has strict preferences while the other side can have both indifferent nodes and nodes with strict preferences. We show that the problem of finding a *strongly* popular matching is polynomial-time solvable even in the latter case. Our result also extends to the many-to-many version, i.e. the strongly popular b-matching problem.

1 Introduction

A **bipartite preference system with ties** consists of a bipartite multigraph $G = (S, T; E)$ and, for every node $v \in S \cup T$, a partial order $\preceq_v$ on the edges incident to v. Given a bipartite preference system with ties, a node **prefers** a matching M_1 to a matching M_2 if it is either matched in M_1 but not in M_2, or matched by a better edge in M_1 than in M_2. Matching M_1 is **more popular** than matching M_2 if the number of nodes preferring M_1 to M_2 is strictly larger than the number of nodes preferring M_2 to M_1. This relation is not transitive; it is possible that M_1 is more popular than M_2, M_2 is more popular than M_3, and M_3 is more popular than M_1 [2]. A matching M is **popular** if no matching is more popular than M, and it is **strongly popular** if M is more popular than any other matching. These notions were first introduced by Gärdenfors [8], who showed that *a)* every strongly popular matching is stable and *b)* in case of no ties, all stable matchings are popular.

Informally, the existence of a strongly popular matching means that the “popularity contest” among matchings has an undisputed winner. There obviously cannot be two distinct strongly popular matchings, because both of them would have to be

[⋆]MTA-ELTE Egerváry Research Group, Department of Operations Research, Eötvös University, Pázmány Péter sétány 1/C, Budapest, Hungary, H-1117. Email: `tkiraly@cs.elte.hu`

[⋆⋆]Department of Operations Research, Eötvös University, Pázmány Péter sétány 1/C, Budapest, Hungary, H-1117. Email: `karkuszsuzsi@gmail.com`

more popular than the other, which is impossible. Furthermore, a strongly popular matching must be a unique popular matching. However, there are instances where the popular matching is unique but it is not strongly popular; we refer the reader to the full version of [2] for an example.

Algorithmic questions about popular matchings have generated a lot of interest lately, see Section 1.1 for a short summary of recent results. Here we just mention that for any preference system with ties (even non-bipartite), it can be decided in polynomial time if a given matching is popular or strongly popular [2]. This means that the decision problem for popular matchings is in the complexity class NP, while the decision problem for strongly popular matchings is in the lesser-known complexity class UP (Unambiguous Polynomial-time). The latter class, introduced by Valiant [17], consists of the decision problems solvable by an NP machine such that all witnesses are rejected in a "no" instance, while exactly one witness is accepted in a "yes" instance. The strongly popular matching problem belongs to this class because each "yes" instance has a single strongly popular matching and it can be verified in polynomial time. No complete problem is known for the class UP, so proving a hardness result for strongly popular matchings would be incredibly challenging.

For a node $v \in S \cup T$, let $\delta_G(v)$ denote the set of edges incident to v in G. We consider bipartite preference systems with two types of nodes: **nodes with strict preferences**, where the preference order $\preceq_v$ is a total order on $\delta_G(v)$, and **indifferent nodes**, where every incident edge is equally good (and better than being unmatched). If all nodes have strict preferences, then every stable matching is popular [8]. On one hand, this implies that there always exists a popular matching and one can be found using the well-known Gale-Shapley algorithm [7]. On the other hand, we can decide if a strongly popular matching exists by finding an arbitrary stable matching and checking whether it is strongly popular (this also works in non-bipartite preference systems without ties [2]).

The problems become more complex when indifferent nodes are also allowed on one of the sides. If nodes on one side have strict preferences while those on the other side are all indifferent, then the existence of a popular matching can still be decided in polynomial time, as shown by Cseh, Huang, and Kavitha [4]. However, they also showed that the problem becomes NP-complete if one side has strict preferences while the other side may feature both indifferent nodes and nodes with strict preferences; see the full version of [4] and [5] for proofs.

In this paper, we prove that the existence of a *strongly* popular matching can be decided in polynomial time even in the latter case.

Theorem 1.1. *Given a bipartite preference system $(G = (S, T; E), \preceq)$ where nodes in S have strict preferences and each node in T is either indifferent or has strict preferences, it can be decided in polynomial time if there is a strongly popular matching.*

Our algorithm successively finds edges that *cannot* be in a strongly popular matching or *must* be in any strongly popular matching, and also maintains a directed graph related to the possible structure of the strongly popular matching. The set of possible candidates keeps shrinking until, at the end, we can either conclude that there is no strongly popular matching, or exactly one candidate matching remains. In the latter case, we can check in polynomial time whether this matching is strongly popular.

Our result also extends to the **strongly popular b-matching problem** in bipartite preference systems. In the b-matching problem, each node v has a quota $b(v) \in \mathbb{Z}_+$. An edge set $M \subseteq E$ is a **b-matching** if $d_M(v) \leq b(v)$ for every v, where $d_M(v)$ denotes the number of edges of M incident to v. There are various ways to define popularity for b-matchings, and in this paper we follow the definition used by Brandl and Kavitha [3] and by Kamiyama [13] (we note that a different definition of popularity is used in [16]). The precise definition of popularity in the many-to-many setting is presented in Section 2, where we also prove the following result, which has not yet been published in the literature.

Theorem 1.2. *In bipartite preference systems with ties, popularity and strong popularity of a given b-matching can be decided in polynomial time.*

As in the matching problem, we consider the setting where one side has strict preferences, while the other side may feature both indifferent nodes and nodes with strict preferences. Indifferent nodes prefer to fill as much of their quota as possible. Such dichotomic behaviour is natural in several real-world allocation problems. For example, in a student allocation problem, some schools may organize exams to rank students, while other schools may not differentiate between admissible students. We prove the following extension of Theorem 1.1.

Theorem 1.3. *Given a bipartite preference system $(G = (S, T; E), \preceq)$ with quotas $b(v) \in \mathbb{Z}_+$ $(v \in S \cup T)$, where nodes in S have strict preferences and each node in T is either indifferent or has strict preferences, it can be decided in polynomial time if there is a strongly popular b-matching.*

The proof is presented in Section 3. Theorem 1.1 is obtained by setting $b \equiv 1$.

1.1 Related work on popular matchings

There is a lot of ongoing research about the computational complexity of the popular matching problem. For bipartite preference systems with no ties, Huang and Kavitha [9] showed that a maximum size popular matching can be found in polynomial time, and Cseh and Kavitha [6] gave an algorithm for deciding if a given edge belongs to a popular matching. On the other hand, the complexity of deciding the existence of a popular matching in a non-bipartite preference system without ties is still open. Huang and Kavitha [10] introduced the notion of unpopularity factor, and showed that, for any positive ε, it is NP-hard to compute a matching with unpopularity factor within $\frac{4}{3} - \varepsilon$ of optimal. In another paper, they showed that the problem of finding a maximum-weight popular matching is NP-hard, while a maximum-weight popular half-integral matching can be found in polynomial time.

Several recent results concern a slightly different, one-sided model (also known as the House Allocation model), where one side has preference lists, while nodes on the other side do not vote at all and do not prefer to be matched. Abraham et al. [1]

gave a polynomial-time algorithm for finding a popular matching in this model. If the preferences are strict, then optimal popular matchings can also be found for various notions of optimality [14, 15].

2 Popular b-matchings

The popular matching problem can be extended to the many-to-one setting (the so-called Hospitals-Residents problem), and also to the many-to-many setting (the bipartite b-matching problem). Two models have been proposed, one by Nasre and Rawat [16], and another by Brandl and Kavitha [3]. In this paper we use the latter model, but introduce it in a slightly different way. Both papers proved that, in case of strict preferences, a maximum size popular matching in the respective models can be found in polynomial time. Kamiyama [13] extended the second model with matroid constraints, and showed how to find a maximum size popular matching if the constraints are defined by weakly base orderable matroids.

Let $G = (S, T; E)$ be a bipartite graph, and let $b : S \cup T \to \mathbb{Z}_+$ denote the quota function. We define the notion of popularity in b-matchings for general bipartite preference systems with ties, where the partial orders $\preceq_v$ $(v \in S \cup T)$ can be arbitrary. For edges $e, f \in \delta_G(v)$, let

$$\mathrm{vote}_v(e, f) = \begin{cases} 1 & \text{if } e \succ_v f \\ -1 & \text{if } e \prec_v f \\ 0 & \text{otherwise.} \end{cases}$$

For technical reasons, we need to allow the empty set as an argument, so we extend the definition by $\mathrm{vote}_v(e, \emptyset) = 1$, $\mathrm{vote}_v(\emptyset, f) = -1$, and $\mathrm{vote}_v(\emptyset, \emptyset) = 0$. Let M_1 and M_2 be b-matchings, and let $v \in S \cup T$. We say that $(e_1, \dots, e_{b(v)}; f_1, \dots, f_{b(v)})$ is a **valid enumeration** of (M_1, M_2) at v if

- each e_i is either an edge in $M_1 \cap \delta_G(v)$ or the empty set
- each f_i is either an edge in $M_2 \cap \delta_G(v)$ or the empty set
- each edge of $M_1 \cap \delta_G(v)$ appears exactly once among the e_is
- each edge of $M_2 \cap \delta_G(v)$ appears exactly once among the f_is
- if $e_i = f_j \in M_1 \cap M_2 \cap \delta_G(v)$, then $i = j$
- if $e_i = \emptyset$ and $f_i \neq \emptyset$ for some i, then there is no j such that $e_i \neq \emptyset$ and $f_i = \emptyset$.

The last property implies that the number of indices i where $e_i \neq \emptyset$ and $f_i \neq \emptyset$ is $\min\{|M_1 \cap \delta_G(v)|, |M_2 \cap \delta_G(v)|\}$. We define

$$\mathrm{vote}_v(M_1, M_2) = \min\{\sum_{i=1}^{b(v)} \mathrm{vote}_v(e_i, f_i) : (e_1, \dots, e_{b(v)}; f_1, \dots, f_{b(v)}) \text{ is a valid enumeration of } (M_1, M_2) \text{ at } v\}.$$

Observe that we take the valid enumeration that is worst from the point of view of M_1. This implies that $\mathrm{vote}_v(M_1, M_2) + \mathrm{vote}_v(M_2, M_1) \le 0$, but equality does not necessarily hold. Define

$$\mathrm{vote}(M_1, M_2) = \sum_{v \in S \cup T} \mathrm{vote}_v(M_1, M_2).$$

A b-matching M is **popular** if $\mathrm{vote}(M, M') \ge 0$ for every b-matching M', and it is **strongly popular** if $\mathrm{vote}(M, M') > 0$ for every b-matching M' distinct from M. If $b \equiv 1$, then these definitions coincide with the standard definitions for matchings. Since $\mathrm{vote}(M_1, M_2) + \mathrm{vote}(M_2, M_1) \le 0$ for any two b-matchings (M_1, M_2), there can be at most one strongly popular b-matching.

In the remainder of the section, we show that popularity and strong popularity of a given b-matching M can be decided in polynomial time. We define an auxiliary bipartite graph $\hat{G}_M = (\hat{S}, \hat{T}; \hat{E}_M)$ where every node $v \in S \cup T$ is replaced by $b(v)$ nodes $\hat{v}_1, \dots, \hat{v}_{b(v)}$. For $st \in E \setminus M$, we introduce edges $\hat{s}_i \hat{t}_j$ for every $i \in [b(s)]$ and $j \in [b(t)]$. We also define edges corresponding to M that form a matching $\hat{M}$: for every edge $st \in M$, we add a single edge $\widehat{st} = \hat{s}_i \hat{t}_j$, in such a way that $\hat{M} = \{\widehat{st} : st \in M\}$ is a matching in $\hat{G}_M$. This is possible because $d_M(v) \le b(v)$ for every $v \in S \cup T$. Thus, an edge $e = st \in E$ has a single corresponding edge in $\hat{E}_M$ if $e \in M$, and $b(s)b(t)$ corresponding edges if $e \in E \setminus M$.

The preference system on G induces a preference system on $\hat{G}_M$, with the additional definition that each node is indifferent between edges in $\hat{E}_M$ corresponding to the same edge in $E \setminus M$. Note that it is *not* true that M is popular if and only if $\hat{M}$ is popular. However, we can characterize the popularity of M using alternating paths with respect to $\hat{M}$ in $\hat{G}_M$, and this leads to a polynomial-time algorithm for deciding if M is popular / strongly popular.

We define a weight function w on $\hat{E}_M$. Let $w(e) = 0$ if $e \in \hat{M}$. For given $\hat{s}_i \hat{t}_j \in \hat{E}_M \setminus \hat{M}$, let e_1 be the edge of $\hat{M}$ incident to $\hat{s}_i$ if $\hat{s}_i$ is covered by $\hat{M}$, otherwise let e_1 be the empty set. Similarly, let e_2 be the edge of $\hat{M}$ incident to $\hat{t}_j$ if $\hat{t}_j$ is covered by $\hat{M}$, otherwise let $e_2 = \emptyset$. Let

$$w(\hat{s}_i \hat{t}_j) = \mathrm{vote}_{\hat{s}_i}(e_1, \hat{s}_i \hat{t}_j) + \mathrm{vote}_{\hat{t}_j}(e_2, \hat{s}_i \hat{t}_j).$$

An **alternating cycle** with respect to $\hat{M}$ is a cycle with edges alternating between $\hat{M}$ and $\hat{E}_M \setminus \hat{M}$. An **alternating path** is a path alternating between $\hat{M}$ and $\hat{E}_M \setminus \hat{M}$, such that if the first or last edge is in $\hat{E}_M \setminus \hat{M}$, then the corresponding end-node of the path is not covered by $\hat{M}$. The **modifier** of an alternating path P, denoted by $\mathrm{mod}(P) \in \{0, 1, 2\}$, is the number of its end-nodes covered by $\hat{M}$. An alternating path P is **invalid** if $\mathrm{mod}(P) = 1$ and its end-nodes are $\hat{v}_i$ and $\hat{v}_j$ for the same v, otherwise it is **valid**.

Theorem 2.1. *A b-matching M is popular if and only if $w(C) \ge 0$ for every alternating cycle C and $w(P) + \mathrm{mod}(P) \ge 0$ for every valid alternating path P in the auxiliary graph $\hat{G}_M$. A b-matching M is strongly popular if and only if $w(C) > 0$ for every alternating cycle C and $w(P) + \mathrm{mod}(P) > 0$ for every valid alternating path P in the auxiliary graph $\hat{G}_M$.*

Proof. We first prove that if M is not popular, then the alternating cycle or path described in the theorem exists (the same proof works for strong popularity). Since M is not popular, there exists a b-matching M' such that $\mathrm{vote}(M, M') < 0$. We may assume that at least one endpoint of every edge of M is covered by M', since otherwise that edge can be added to M'.

We define a matching $\hat{M}'$ in $\hat{G}_M$ based on the voting at the individual nodes. If $st \in M \cap M'$, then we include $\widehat{st}$ in $\hat{M}'$. For every $v \in S \cup T$, we fix a valid enumeration $(e_1^v, \dots, e_{b(v)}^v; f_1^v, \dots, f_{b(v)}^v)$ of (M, M') at v that is a minimizer of $\mathrm{vote}_v(M, M')$ and satisfies the following property: if e_i^v is an edge vw of M, then $\widehat{vw} = \hat{v}_i \hat{w}_j$ for some j. It is easy to see that there is an enumeration with this property, since we can permute the indices arbitrarily. If $f = st$ is an edge in $M' \setminus M$, then there are indices i and j such that $f = f_i^s$ and $f = f_j^t$. We include $\hat{s}_i \hat{t}_j$ in $\hat{M}'$, and denote it by $\hat{f}$.

Consider $\hat{M} \Delta \hat{M}'$, i.e. the symmetric difference of $\hat{M}$ and $\hat{M}'$. Since its degrees are at most 2, $\hat{M} \Delta \hat{M}'$ is the disjoint union of cycles $C_1, \dots, C_k$ and paths $P_1, \dots, P_l$. Observe that $C_1, \dots, C_k$ are alternating cycles and $P_1, \dots, P_l$ are alternating paths. Furthermore, each P_i is valid, because the validity of the enumeration at v means that if some v_i is covered by $\hat{M}'$ but not by $\hat{M}$, then any v_j covered by $\hat{M}$ is also covered by $\hat{M}'$.

Claim 2.2. $\mathrm{vote}(M, M') = \sum_{i=1}^k w(C_i) + \sum_{j=1}^l (w(P_j) + \mathrm{mod}(P_j))$.

Proof. Consider an edge $f = st$ in $M' \setminus M$, and the corresponding edge $\hat{f} = \hat{s}_i \hat{t}_j$ in $\hat{M}'$. Let e_1 (e_2) be the edge of $\hat{M}$ incident to $\hat{s}_i$ ($\hat{t}_j$) if it is covered by $\hat{M}$, and the empty set otherwise. By definition, $f = f_i^s$ and $f = f_j^t$, so

$$w(\hat{f}) = \mathrm{vote}_{\hat{s}_i}(e_1, \hat{f}) + \mathrm{vote}_{\hat{t}_j}(e_2, \hat{f}) = \mathrm{vote}_s(e_i^s, f_i^s) + \mathrm{vote}_t(e_j^t, f_j^t).$$

This means that $\sum_{i=1}^k w(C_i) + \sum_{j=1}^l w(P_j)$ counts all votes except those of type $\mathrm{vote}_v(e_i^v, f_i^v)$ where $e_i^v \in M$ and $f_i^v = \emptyset$. The number of these votes is exactly the number of nodes covered by $\hat{M}$ but not by $\hat{M}'$, which equals $\sum_{j=1}^l \mathrm{mod}(P_j)$. □

We obtained that $\sum_{i=1}^k w(C_i) + \sum_{j=1}^l (w(P_j) + \mathrm{mod}(P_j)) = \mathrm{vote}(M, M') < 0$, so either there is a cycle C_i for which $w(C_i) < 0$, or there is a path P_j for which $w(P_j) + \mathrm{mod}(P_j) < 0$. This proves the "if" direction of the theorem.

To prove the "only if" direction, we examine two cases.

Case 1: there is an alternating cycle C with $w(C) < 0$. We cannot directly use the cycle to construct a b-matching M' and a valid enumeration of (M, M') that shows $\mathrm{vote}(M, M') < 0$, because the enumeration determined by the cycle may violate the last property in the definition of valid enumerations. However, this may happen only if C contains distinct edges $\hat{s}_{i_1} \hat{t}_{j_1}$ and $\hat{s}_{i_2} \hat{t}_{j_2}$ for some $st \in E \setminus M$. The following claim shows how to avoid this problem.

Claim 2.3. *If an alternating cycle C contains distinct edges $\hat{s}_{i_1} \hat{t}_{j_1}$ and $\hat{s}_{i_2} \hat{t}_{j_2}$ corresponding to the same edge $st \in E \setminus M$, then, by removing these edges and adding $\hat{s}_{i_1} \hat{t}_{j_2}$ and $\hat{s}_{i_2} \hat{t}_{j_1}$, we get two alternating cycles C_1, C_2 with $w(C_1) + w(C_2) = w(C)$.*

Proof. The exchange results in 2 cycles because the graph is bipartite. The definition of w implies that $w(\hat{s}_{i_1}\hat{t}_{j_1}) + w(\hat{s}_{i_2}\hat{t}_{j_2}) = w(\hat{s}_{i_1}\hat{t}_{j_2}) + w(\hat{s}_{i_2}\hat{t}_{j_1})$, so $w(C_1) + w(C_2) = w(C)$. $\square$

Since $w(C) < 0$, one of the resulting cycles has negative weight. We can repeat similar operations until we get a cycle C' with $w(C') < 0$ which in addition does not contain distinct edges corresponding to the same edge of $E \setminus M$. By exchanging along the cycle, we obtain a matching $\hat{M}'$, which determines a b-matching M' and a valid enumeration of (M, M') at each node. This implies $\text{vote}(M, M') \leq w(C') < 0$.

Case 2: there is a valid alternating path P with $w(P) + \text{mod}(P) < 0$. As in Case 1, some modifications are needed in order to obtain valid enumerations. The counterpart of Claim 2.3 for this case is the following.

Claim 2.4. *If a valid alternating path P contains distinct edges $\hat{s}_{i_1}\hat{t}_{j_1}$ and $\hat{s}_{i_2}\hat{t}_{j_2}$ corresponding to the same edge $st \in E \setminus M$, then, by removing these edges and adding $\hat{s}_{i_1}\hat{t}_{j_2}$ and $\hat{s}_{i_2}\hat{t}_{j_1}$, we get an alternating cycle C' and a valid alternating path P' with $w(C') + w(P') = w(P)$ and $\text{mod}(P') = \text{mod}(P)$.*

Proof. The proof is the same as the proof of Claim 2.3, with the additional observation that P' has the same end-nodes as P, so P' is also valid and $\text{mod}(P') = \text{mod}(P)$. $\square$

The claim implies that, by repeating this operation as many times as needed, we either get a a cycle C' with $w(C') < 0$ or a valid path P' with $w(P') + \text{mod}(P') < 0$, having the additional property that it does not use distinct edges corresponding to the same edge of $E \setminus M$. By exchanging along the path or cycle, we get a matching $\hat{M}'$, which determines a b-matching M' and a valid enumeration of (M, M') at each node. In the cycle case we have $\text{vote}(M, M') \leq w(C') < 0$. In the path case, voting according to this valid enumeration gives the result $w(P') + \text{mod}(P')$, because the number of nodes covered by $\hat{M}$ but not by $\hat{M}'$ is $\text{mod}(P')$. This implies $\text{vote}(M, M') \leq w(P') + \text{mod}(P') < 0$. This concludes the proof for popularity. The proof for strong popularity is analogous, with $\text{vote}(M, M') \leq 0$ in place of $\text{vote}(M, M') < 0$. $\square$

Using Theorem 2.1, it is straightforward to check the popularity or strong popularity of a b-matching M.

Proof of Theorem 1.2. Given a b-matching M, we construct the auxiliary bipartite graph $\hat{G}_M$ and the weight function w as specified in the first part of the section. Let D_M be the directed graph obtained from $\hat{G}_M$ by orienting the edges of $\hat{M}$ from $\hat{S}$ to $\hat{T}$ and the edges of $\hat{E} \setminus \hat{M}$ from $\hat{T}$ to $\hat{S}$. We can use the Bellman-Ford algorithm to check if there is a directed cycle C with $w(C) < 0$ or $w(C) = 0$. If there is no negative cycle, then we can compute the minimum weight paths between all pairs of nodes. Based on this, we can decide if there is a valid alternating path P with $w(P) + \text{mod}(P) < 0$ or $w(P) + \text{mod}(P) = 0$. $\square$

3 Proof of the main theorem

In this section we prove Theorem 1.3. We are given a bipartite multigraph $G = (S, T; E)$ and a quota function $b : S \cup T \to \mathbb{Z}_+$. The node set T is partitioned into two parts, T_P and T_I. Every node $v \in S \cup T_P$ has a strict preference order $\preceq_v$ over its incident edges, while the nodes in T_I are indifferent but prefer to fill their quotas. We give a polynomial-time algorithm which decides if the instance admits a strongly popular b-matching (SPbM for short), as defined in Section 2.

3.1 Preliminaries

Before going into the details, we give an overview of the main ideas of the proof. We may assume, without loss of generality, that G has no isolated nodes and that $b(v) \le d_G(v)$ for every $v \in S \cup T$. During the algorithm, we modify the instance using the following two **reduction operations**.

1. We remove edges that cannot appear in an SPbM of the current instance. We remove any isolated nodes that arise, and we also decrease quotas if they exceed the degree.

2. We fix edges that are guaranteed to belong to the SPbM of the current instance if it has one. Fixed edges are removed, and the quotas of their two endnodes are decreased by one. Nodes that become isolated are also removed.

Let $G^k = (S^k, T^k; E^k)$ be the instance after performing k of the above reduction operations, let b^k be the corresponding quota function, and let F be the set of edges fixed so far.

Lemma 3.1. *If the original instance has an SPbM M, then $F \subseteq M$, and $M \setminus F$ is an SPb^kM in G^k.*

Proof. The proof is by induction on k; the claim is obviously true for $k = 0$. Let G^{k-1} be the instance before the last operation. If the last operation was the removal of an edge st, then, by induction, M contains F, $M \setminus F$ is an SPb^{k-1}M of G^{k-1}, and $st \notin M$. Thus $M \setminus F$ is an SPb^kM of G^k.

If we fixed an edge st in the last operation, then $M \setminus (F - st)$ is an SPb^{k-1}M of G^{k-1} by induction, and $st \in M \setminus (F - st)$ because we only fix edges that are guaranteed to belong to the SPb^{k-1}M. This implies that $st \in M$. Since b^k is obtained from b^{k-1} by decreasing the quota of s and t by one, $M \setminus F$ is an SPb^kM of G^k. $\square$

Note that it is possible that G^k has an SPb^kM even though G does not have an SPbM. This is not a problem however: if we eventually obtain an empty graph by reduction operations, then F is the only candidate for an SPbM by Lemma 3.1, and we can check in polynomial time if it is an SPbM of G or not. On the other hand, if we obtain a graph G^k that has no SPb^kM, then G has no SPbM by Lemma 3.1.

To summarize, our strategy is to do reduction operations until we get to the empty graph or we get a certificate that there is no SPbM. The question is how to identify edges that can be removed or fixed. In the proof, we will present a sequence of *claims* about certain edges being removable or fixable. Each claim is constructive in the sense that the specified edges can be found in linear time.

Each claim is followed by the description of a *property* satisfied by any instance that cannot be further reduced by the reduction operations described in the claim. At any point in the proof, we assume (without explicitly stating it) that our instance G has all the properties described previously. This assumption is valid because the algorithm can make reductions until all properties are satisfied.

Additional definitions. An edge $st \in E$ is called a **blocking edge** with respect to a b-matching M if it satisfies the following three conditions: *i)* $st \notin M$, *ii)* s does not fill its quota in M or there is a node t' such that $st' \in M$ and $t \succ_s t'$, and *iii)* t does not fill its quota in M or there is a node s' such that $s't \in M$ and $s \succ_t s'$. If M is an SPbM, then there is no blocking edge with respect to M; indeed, if M' is the b-matching obtained from M by adding a blocking edge st and removing st' if s fills its quota in M and removing $s't$ if t fills its quota in M, then vote$(M, M') \leq 0$, which contradicts that M is an SPbM.

In addition to blocking edges, we will use *alternating paths and cycles* to show that certain b-matchings cannot be strongly popular. Note that these are paths and cycles in G, not in the auxiliary graph $\hat{G}_M$ defined in Section 2 (we do not use $\hat{G}_M$ in this section). Given a b-matching M and an alternating path or cycle w.r.t. M, let M' be the b-matching obtained from M by exchanging along the path or cycle – if we exchange along a path whose first or last edge is not in M and the endpoint fills its quota, then we also remove the worst edge of M covering the corresponding endpoint of the path (if the node is indifferent, then we remove an arbitrary incident edge in M). If we can show that vote$(M, M') \leq 0$, then M is not an SPbM.

At some points in the proof (proofs of Claim 3.6, Lemmas 3.9 and 3.10), we append an extra edge ts' to an alternating path ending in $st \in M$. We call the result an **alternating quasi-path** if s' appears earlier on the path. If we exchange along the quasi-path and s' fills its quota in M, then, as above, we have to remove the worst edge $s't'$ of M at s'. This is a problem if $s't'$ belongs to the quasi-path, because we cannot remove it twice. However, in the situations where we use this technique, ts' is guaranteed to be better than $s't'$ at s', and if $s't'$ is in the quasi-path, then we can exchange along the cycle part of the quasi-path to get a b-matching M' with vote$(M, M') \leq 0$.

The following *auxiliary digraph* D plays an important role in the proof: for every node $v \in S \cup T_P$, we orient the first $b(v)$ edges on v's preference list towards their other endpoint. Some edges may be bidirected; however, as our first claim (Claim 3.2) will show, bidirected edges can always be fixed. The digraph D has to be recomputed after each reduction.

3.2 Description of the algorithm

We assume that G has no isolated nodes and that $b(v) \leq d_G(v)$ for every $v \in S \cup T$.

Orientation. For every node $v \in S \cup T_P$, we orient the first $b(v)$ edges on v's preference list towards their other endpoint. The digraph obtained this way is denoted by D, and it is recomputed after every reduction.

Claim 3.2. *If an edge $st \in E$ is bidirected in D, then it belongs to the SPbM if there is one.*

Proof. If M is a b-matching and $st \notin M$, then st is blocking with respect to M. □

Property 1. *After reductions according to Claim 3.2, D does not contain bidirected edges.*

Claim 3.3. *If there are $b(v)$ directed edges to a node $v \in S \cup T_P$: $w_1v, w_2v, \dots, v_{b(v)}v$, and uv is an edge such that $uv \prec_v w_iv$ for every $1 \leq i \leq b(v)$, then uv cannot belong to the SPbM.*

Proof. Suppose that M is an SPbM and $uv \in M$. Then there is an index $1 \leq i \leq b(v)$ such that $w_iv \notin M$, but then w_iv is a blocking edge with respect to M. □

Property 2. *After reductions according to the previous claims, every node $v \in S \cup T_P$ is entered by at most $b(v)$ directed edges in D.*

Claim 3.4. *If there are at most $b(t)$ directed edges entering a node $t \in T_I$, then these edges belong to the SPbM if there is one.*

Proof. Suppose that the SPbM M does not contain one of these edges st; then t has to fill its quota, otherwise st would be blocking. Since there are at most $b(t)$ directed edges entering t, there has to be an edge ut that belongs to M and is not directed towards t. There are $b(u)$ directed edges from u, therefore there is a directed edge ut_2 which does not belong to M. Suppose that $t_2 \in T_P$. There are at most $b(t_2)$ directed edges entering t_2 (because of Claim 3.3), and t_2 fills its quota (otherwise ut_2 would be blocking), therefore there is an edge $u_2t_2 \in M$ such that it is not directed towards t_2. Consider the path that starts with s and alternates between directed edges that do not belong to M and edges of M that are not directed backwards in the path. (The first edge is st, the second is tu, the third is ut_2.) If we reach a node $t' \in T_I$, then by exchanging along the path we get a matching M' such that vote$(M, M') \leq 0$. Indeed, the vote of the nodes of S on the path is -1, while only the following nodes may have a positive vote: the nodes of T in the path except for t and t', the other endpoint of the edge of M incident to t' (if it exists), and the other endpoint of the edge of M incident to s. This contradicts the assumption that M is an SPbM.

If at some point the path returns to a previous node, then we get an alternating cycle, and exchanging along the cycle yields a matching M' such that vote$(M, M') \leq 0$. See Figure 1 for an illustration of both cases. □

Property 3. *After reductions according to the previous claims, each node $t \in T_I$ is entered by either 0 or more than $b(t)$ directed edges in D.*

Definition. Let T_1 be the set of nodes of T_I with in-degree more than $b(t)$ in D, and let T_2 be the set of nodes of T_I with no incoming directed edges.

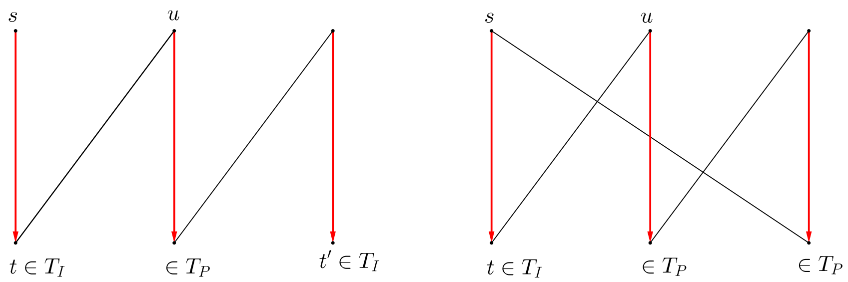

Figure 1: The black edges belong to the SPbM.

Claim 3.5. *Suppose that the number of directed edges entering a node $s \in S$ plus the number of edges between s and T_2 is more than $b(s)$. Let sv be the least preferred by s among these edges. Then sv cannot belong to the SPbM.*

Proof. Suppose that sv belongs to the SPbM M. Every directed edge entering s has to belong to M too, because these are preferred over sv and therefore they would be blocking. Thus, there has to be an edge $st \notin M$ with $t \in T_2$. There is an edge $ts' \in M$ because t fills its quota (otherwise st would be blocking). There are $b(s')$ directed edges leaving s', therefore there is one which is not in M. Consider the alternating path that starts with vs, st, ts' and then alternates between directed edges that do not belong to M and edges of M that are not directed backwards in the path. If at some point this path returns to a previous node of the path, then we get an alternating cycle, otherwise we reach a node in T_1. Let M' be the b-matching we get from M by exchanging along the cycle or path; this satisfies $\text{vote}(M, M') \leq 0$. ☐

Property 4. *After reductions according to the previous claims, the number of directed edges entering a node $s \in S$ plus the number of edges between s and T_2 is at most $b(s)$.*

Claim 3.6. *If uv is an edge in $G[S \cup T_P \cup T_1]$ and it is not a directed edge in any direction, then uv cannot belong to the SPbM.*

Proof. Suppose that uv is in the SPbM M. If $u \in T_1$, then there are more than $b(u)$ directed edges entering u. If such an edge has tail v, then it is also in M, otherwise it would be blocking. Therefore there is a directed edge su for some $s \neq v$ which is not in M. There are $b(v)$ directed edges leaving v, therefore there is a directed edge $vt \notin M$. If $t \in T_p$, then there are at most $b(t)$ directed edges entering t, therefore there is an edge $s't \in M$ which is not directed towards t since t has to fill its quota (otherwise vt would be blocking). Consider the path starting with su, uv and then alternating between directed edges that are not in M and edges of M that are not directed backwards in the path. Similarly to the proof of Claim 3.4, we either reach a node in T_I or return to a previous node of the path, and exchanging along the obtained path or cycle yields a matching M' such that $\text{vote}(M, M') \leq 0$.

Now consider the case where $u \in T_P$. First we build a path starting with vu and then alternating between directed edges that are not in M and edges of M in

$G[S \cup T_P \cup T_1]$ that are not directed backwards. The latter can be achieved because a node $s \in S$ entered by a directed edge not in M fills its quota, so by Property 4 one of its neighbours in M is not in T_2. If we return to a previous node of the path without reaching a node in T_1, then exchanging along the cycle yields a matching M' and $\text{vote}(M, M') \leq 0$. If we reach a node in T_1, then there is a directed edge not in M pointing to this node, which we add to the path. Let this (quasi-)path be denoted by P. We continue P from v with edges alternating between directed edges not in M and edges of M that are not directed backwards. If we reach a node in T_1, then exchanging along the (quasi-)path yields a matching M' and $vote(M, M') \leq 0$; see Figure 2 for an illustration of this case. If we return to a previous node of the path, then, again, exchanging along the obtained cycle yields a matching M' and $vote(M, M') \leq 0$. $\square$

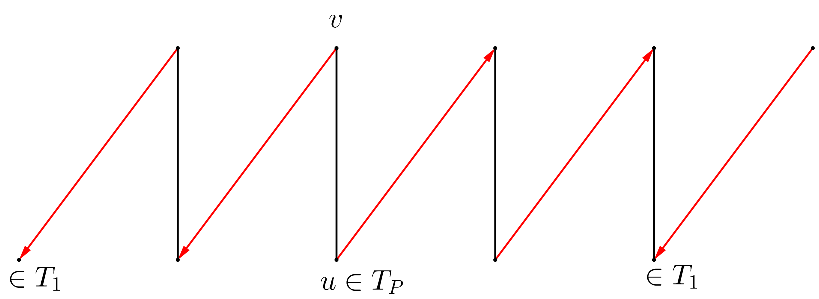

Figure 2: The black edges belong to the SPbM.

Property 5. *After reductions according to the previous claims, every edge of G is either directed in D or has one endpoint in T_2.*

Lemma 3.7. *Suppose there is an SPbM M, and a directed edge $st \notin M$. Then there is a directed path from s to a node in T_1 that starts with st and which is alternating with respect to M.*

Proof. We claim that if $t \in T_P$, then there is a directed edge ts_2 in M. Indeed, t fills its quota (otherwise st would be blocking), there are at most $b(t)$ directed edges entering t, and only edges of D are incident to t; therefore, there is at least one outgoing directed edge ts_2 that belongs to M.

There are $b(s_2)$ directed edges leaving s_2, so there is a directed edge $s_2t_2 \notin M$. Continuing the alternating directed path like this we either reach a node in T_1 or we get an alternating directed cycle which contradicts the strong popularity of M. $\square$

Lemma 3.8. *Suppose there is an SPbM M, and a directed edge $st \in M$. Then there is a directed path from s to a node in T_1 that starts with st and which is alternating with respect to M.*

Proof. If $t \in T_P$, then there is a directed edge $ts_2 \notin M$, because there are $b(t)$ directed edges leaving t. The node s_2 fills its quota since otherwise ts_2 would be blocking; therefore, there is a directed edge $s_2t_2 \in M$ because of Property 4. Continuing the alternating directed path like this we either reach a node in T_1 or we get an alternating directed cycle which contradicts that M is strongly popular. □

Lemma 3.9. *If M is an SPbM, then $d_M(v) = b(v)$ for every $v \in S \cup T$.*

Proof. Let M be an SPbM, and suppose that a node $s \in S$ does not fill its quota. Then there is a directed edge leaving s that is not in M. By Lemma 3.7, there is an alternating directed path from s to T_1 that starts with this edge. Exchanging along this alternating path yields a b-matching M' with $\text{vote}(M, M') \leq 0$.

Now suppose that a node $t \in T_P$ does not fill its quota. There is a directed edge $ts \notin M$, and s fills its quota, otherwise ts would be blocking. By Property 4, there is a directed edge $st' \in M$ with $t' \neq t$, and by Lemma 3.8 there is an alternating directed path P from s to a node $t_1 \in T_1$ that starts with st'. If this path reaches t, then we get an alternating directed cycle which contradicts the strong popularity of M. If not, let $s't_1$ be a directed edge which does not belong to M (such an edge exists by Property 3). Let M' be the b-matching we get from M by exchanging along the alternating (quasi-)path $ts \cup P \cup t_1 s'$. This satisfies $\text{vote}(M, M') \leq 0$.

A node $t \in T_1$ fills its quota since there are more than $b(t)$ directed edges entering t by Property 3.

Finally, suppose that a node $t \in T_2$ does not fill its quota. Let st be an edge that is not in M. It follows from Property 4 that there is a directed edge $st' \in M$ since s fills its quota. By Lemma 3.8, there is a directed alternating path P from s to a node $t_1 \in T_1$ that starts with st'. Let $s't_1$ be a directed edge that does not belong to M. Let M' be the b-matching we get from M by exchanging along the alternating (quasi-)path $ts \cup P \cup t_1 s'$. This satisfies $\text{vote}(M, M') \leq 0$. □

Lemma 3.10. *Let M be an SPbM and $st \in M$ a directed edge. If s prefers st' over st, then $st' \in M$.*

Proof. Suppose $st' \notin M$. By Lemma 3.8, there is an alternating directed path P from s to a node $t_1 \in T_1$ which starts with st, and by Lemma 3.7 there is an alternating directed path P' from s to a node $t_1' \in T_1$ which starts with st'. If P' intersects P, then we get an alternating cycle, and exchanging along this cycle yields a b-matching M' such that $\text{vote}(M, M') \leq 0$. Otherwise, let $s't_1$ be a directed edge that does not belong to M, and let M' be the b-matching we get from M by exchanging along the alternating (quasi-)path $s't_1 \cup P \cup P'$. This satisfies $\text{vote}(M, M') \leq 0$. □

Claim 3.11. *If the number of directed edges entering a node $s \in S$ plus the number of edges between s and T_2 are less than $b(s)$, then s's most preferred edge has to belong to the SPbM if there is one.*

Proof. This follows from Lemmas 3.9 and 3.10. □

Property 6. *After reductions according to the previous claims, the number of directed edges entering a node $s \in S$ plus the number of edges between s and T_2 is exactly $b(s)$.*

Lemma 3.12. *If G is not the empty graph (and all the listed properties hold), then there is no SPbM.*

Proof. Suppose G is not empty and M is an SPbM. If a directed edge st is in M, then there is either a directed edge entering s that does not belong to M, or there is an edge $st' \notin M$ with $t' \in T_2$ because of Property 6 and Lemma 3.9. If a directed edge ts is not in M, then there is a directed edge in M entering t since t fills its quota.

It follows that given a directed edge $st \in M$, we can go on a backwards directed alternating path from s until we reach a node s' such that there is an edge $s'u \notin M$ with $u \in T_2$ (the backwards walk cannot intersect itself because that would give a directed alternating cycle showing that M is not strongly popular). Thus, for every directed edge $st \in M$ with $t \in T_1$, there is a node $s' \in S$ such that there is a directed alternating path P from s' to t with last edge st, and there is an edge $s'u \notin M$ with $u \in T_2$. We call the path $us \cup P$ a **type 1 path** from t to u. From every node in T_1 there is a type 1 path to a node in T_2.

Let us be an edge in M with $u \in T_2$. There is a directed edge $st \notin M$. By Lemma 3.7, there is a directed alternating path P from s to a node $t' \in T_1$ which starts with the edge st. We call the path $us \cup P$ a **type 2 path** from u to t'. From every node in T_2 there is a type 2 path to a node in T_1.

If we go back and forth between T_1 and T_2 on type 1 and type 2 paths, at some point we will return to a node in T_1 or T_2 that we have already touched. We may choose these paths to be edge-disjoint (although not necessarily node-disjoint), because when we reach a node in $S \cup T_P$ that we have already used, we can always choose to proceed on an edge not yet used. Indeed, if there are ℓ non-outgoing edges in M incident to a node $s \in S$, then there are at least ℓ edges not in M leaving s, and if there are ℓ edges in M leaving s, then there are at least ℓ non-outgoing edges not in M incident to s (because of Lemma 3.9 and Property 6). Similarly, if there are ℓ edges not in M entering a node $t \in T_P$, then there are at least ℓ edges in M leaving t, and if there are ℓ edges not in M leaving t, then there are at least ℓ edges in M entering t (because of Lemma 3.9 and Property 2).

We can conclude that if G is not the empty graph, then there are nodes $t_1, t_2, \dots, t_k \in T_1$ and $u_1, u_2, \dots, u_k \in T_2$ such that there is a type 1 path P_i from t_i to u_i and a type 2 path P'_i from u_i to t_{i+1} for $i = 1, \dots, k$ where $t_{k+1} := t_1$, and all these paths are edge-disjoint.

Let C denote the closed tour $P_1 \cup P'_1 \cup P_2 \cup P'_2 \cup \dots \cup P_k \cup P'_k$. The nodes of C in order are $v_1 v_2 \dots v_n$, where $v_1 = v_n$. These nodes are not necessarily distinct; let $V(C)$ denote the node set of C without multiplicities. The tour C is alternating with respect to M. Let M' be the b-matching we get from M by exchanging the edges along C.

For each v_j, let $(e^j_1, \dots, e^j_{b(v_j)}; f^j_1, \dots, f^j_{b(v_j)})$ be the valid enumeration of (M, M') at v_j where e^j_i and f^j_i are consecutive edges of C whenever $e^j_i \neq f^j_i$. Then $\mathrm{vote}_{v_j}(M, M') \leq \sum_{i=1}^{b(v_j)} \mathrm{vote}_{v_j}(e^j_i, f^j_i)$. Observe that $v_i v_{i+1} \in M$ if i is odd and $v_i v_{i+1} \in M'$ if i is even.

Therefore

$$\begin{aligned}\text{vote}(M, M') &= \sum_{v_j \in V(C)} \text{vote}_{v_j}(M, M') \\ &\leq \sum_{j \text{ is odd}} \text{vote}_{v_j}(v_j v_{j+1}, v_{j-1} v_j) + \sum_{j \text{ is even}} \text{vote}_{v_j}(v_{j-1} v_j, v_j v_{j+1}).\end{aligned}$$

The right hand side is equal to 0, because

- if we sum up the votes of the nodes in $T_P \cup S$ of a type 1 path between the two edges of the path incident to that node, we get $+1$
- if we sum up the votes of the nodes in $T_P \cup S$ of a type 2 path between the two edges of the path incident to that node, we get -1
- the vote of the nodes in $T_1 \cup T_2$ is 0.

We obtained that $\text{vote}(M, M') \leq 0$, thus M is not strongly popular. □

Lemma 3.12 shows that by performing the reduction operations defined in the claims, we either get a certificate that the graph has no SPbM (which implies that the original graph also has none), or we eventually reach the empty graph. The latter means that the only possible candidate for an SPbM in the original graph is the set of fixed edges F. We can check in polynomial time if F is an SPbM of the original preference system or not.

4 Conclusion

We proved that in case of strict preferences on one side and both strict preferences and indifference on the other side, the existence of a strongly popular b-matching can be decided in polynomial time. This is a clear indication that the strongly popular matching problem is significantly easier than the popular matching problem. It seems to be difficult to complement this with hardness results; as mentioned in the introduction, the strongly popular matching problem is in the complexity class UP, for which no complete problems are known. Therefore, the more promising direction is to attempt to show polynomial-time solvability for other types of preference systems. In particular, the decision problem for strongly popular matchings is open in the following two cases:

- bipartite preference systems with strict preference and indifference allowed on both sides,
- bipartite preference systems with strict preferences on one side, and arbitrary preferences on the other side.

Our techniques do not seem to extend easily to these problems.

Acknowledgement

The research was supported by the Hungarian National Research, Development and Innovation Office – NKFIH, grants K109240 and K120254.

References

[1] D.J. Abraham, R.W. Irving, T. Kavitha, K. Mehlhorn, *Popular matchings*, SIAM Journal on Computing 37 (2007), 1030–1045.

[2] P. Biró, R. W. Irving, and D. F. Manlove, *Popular matchings in the marriage and roommates problems*, In Proceedings of the 7th International Conference on Algorithms and Complexity, volume 6078 of Lecture Notes in Computer Science (2010), 97–108. Full version: Technical Report TR-2009-306, University of Glasgow, Department of Computing Science.

[3] F. Brandl, T. Kavitha, *Popular Matchings with Multiple Partners*, arXiv: 1609.07531 (2016).

[4] Á. Cseh, C.-C. Huang, and T. Kavitha, *Popular matchings with two-sided preferences and one-sided ties*, In M. M. Halldórsson, K. Iwama, N. Kobayashi, and B. Speckmann, editors, Automata, Languages, and Programming, volume 9134 of Lecture Notes in Computer Science (2015), 367–379. Full version: arXiv: 1603.07168

[5] Á. Cseh, *Complexity and algorithms in matching problems under preferences*, Ph.D. Thesis, TU Berlin (2016).

[6] Á. Cseh, T. Kavitha, *Popular Edges and Dominant Matchings*, Integer Programming and Combinatorial Optimization, Volume 9682 of the series Lecture Notes in Computer Science (2016), 138–151.

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

[8] P. Gärdenfors, *Match making: assignments based on bilateral preferences*, Behavioural Science 20 (1975), 166–173.

[9] C.-C. Huang, T. Kavitha, *Popular Matchings in the Stable Marriage Problem*, Automata, Languages and Programming, Volume 6755 of the series Lecture Notes in Computer Science (2011), 666–677.

[10] C.-C. Huang, T. Kavitha, *Near-Popular Matchings in the Roommates Problem*, SIAM J. Discrete Math. 27 (2013), 43–62.

[11] C.-C. Huang, T. Kavitha, *Popularity, Mixed Matchings, and Self-duality*, Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (2017), 2294–2310.

[12] R.W. Irving, *An efficient algorithm for the "stable roommates" problem*, Journal of Algorithms 6 (1985), 577–595.

[13] N. Kamiyama, *Popular Matchings with Two-Sided Preference Lists and Matroid Constraints*, MI Preprint Series 2016-13, Kyushu University Fukuoka, Japan

[14] T. Kavitha, M. Nasre, *Optimal popular matchings*, Discrete Applied Mathematics 157 (2009), 3181–3186.

[15] E. McDermid, R.W. Irving, *Popular matchings: Structure and algorithms*, In Proceedings of COCOON 2009: the 15th Annual International Computing and Combinatorics Conference, volume 5609 of Lecture Notes in Computer Science (2009), 506–515.

[16] M. Nasre, A. Rawat, *Popularity in the generalized Hospital Residents Setting*, arXiv:1609.07650 (2016).

[17] L. G. Valiant, *Relative complexity of checking and evaluating*, Information Processing Letters 5 (1976), 20–23.