It Only Takes a Few: On the Hardness of Voting With a Constant Number of Agents

Felix Brandt
Institut für Informatik
TU München
brandtf@in.tum.de

Paul Harrenstein
Dept. of Computer Science
University of Oxford
paulhar@cs.ox.ac.uk

Keyvan Kardel
Institut für Informatik
TU München
k.kardel@tum.de

Hans Georg Seedig
Institut für Informatik
TU München
seedigh@in.tum.de

ABSTRACT

Many hardness results in computational social choice make use of the fact that every directed graph may be induced by the pairwise majority relation. However, this fact requires that the number of voters is almost linear in the number of alternatives. It is therefore unclear whether existing hardness results remain intact when the number of voters is bounded, as is for example typically the case in search engine aggregation settings. In this paper, we provide sufficient conditions for majority graphs to be obtainable using a constant number of voters and leverage these conditions to show that winner determination for the Banks set, the tournament equilibrium set, Slater's rule, and ranked pairs remains hard even when there is only a small constant number of voters.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity; J.4 [**Computer Applications**]: Social and Behavioral Sciences - Economics

General Terms

Algorithms, Economics, Theory

Keywords

Social choice and voting, computational complexity

1. INTRODUCTION

A large part of *computational social choice*, a new interdisciplinary area of study at the intersection of social choice theory and computer science, is concerned with the computational complexity of voting problems. For most of the voting rules proposed in the social choice literature, it has been studied how hard it is to determine winners, to identify beneficial strategic manipulations, or to influence the outcome by bribing, partitioning, adding, or deleting voters (see, e.g., [4, 10]). In many cases, the corresponding problems turned out to be NP-hard. Depending on the nature of the problem, this can be interpreted as bad news—as in the case of winner determination—or good news—as in the case of manipulation, bribery, and control.

A large number of voting rules is based on the pairwise majority relation, which establishes a very useful connection between voting theory and graph theory. Perhaps the most fundamental result in this context is *McGarvey's theorem*, which states that *every* directed graph may be induced by the pairwise majority relation [14]. Unsurprisingly, McGarvey's theorem is the basis for most hardness results concerning majoritarian voting rules.[1] McGarvey's original construction requires $2n(n-1)$ voters, where n is the number of alternatives. This number has subsequently been improved by Stearns [17] and Erdős and Moser [9], who have eventually shown that the number of required voters is of order $\Theta(n/\log n)$. As a consequence, the mentioned hardness results only hold if the number of voters is roughly of the same order as the number of alternatives. In many applications, however, the number of voters is much smaller than the number of alternatives and it is unclear whether hardness still holds. A typical example is search engine aggregation, where the voters correspond to Internet search engines and the alternatives correspond to the webpages ranked by the search engines (see, e.g., [8]). Hudry [11] writes that "to my knowledge, when not trivial, the complexity for lower values of m [the number of voters] remains unknown. In particular, it would be interesting to know whether some of the problems [...] remain NP-hard if m is a given constant."

In this paper, we analyze the structure of majority graphs obtainable using a constant number of voters. Obviously, the less voters there are, the more restricted is the corresponding class of inducible majority graphs. For instance, graphs induced by two voters have to be acyclic (and are subject to some additional restrictions). After introducing some basic notation and terminology in Section 2, we completely characterize graphs inducible by two and three voters, respectively, and provide sufficient conditions for graphs to be

[1]A notable exception is a proof by Dwork et al. [8] showing the hardness of Kemeny's rule for any even number of voters greater than two.

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013)*, Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.

induced by k voters in Section 3. In Sections 4, 5, 6, and 7, we leverage these conditions to investigate whether common, computationally intractable voting rules (the Banks set, the tournament equilibrium set, Slater's rule, and ranked pairs) remain intractable when there is only a small constant number of voters. This is achieved by analyzing existing hardness proofs and checking whether the class of majority graphs used in these constructions can be induced by small constant numbers of voters. Somewhat surprisingly, it turns out that all hardness proofs we studied can indeed be constructed using few voters. The paper concludes with an overview of our results, summarized in Table 1, and a brief outlook on future research in Section 8.

2. PRELIMINARIES

This section contains the notation and terminology required to state our results.

A *directed graph* or *digraph* is a pair (V, E), where V is a set of vertices and $E \subseteq V \times V$. As a useful notational convention we adopt $[E] = E \cup \overline{E}$, where $\overline{E} = \{(w, v): (v, w) \in E\}$, i.e., the *converse of* E. We say that E_1 and E_2 are *orientation compatible* if $E_1 \cap ([E_1] \cap [E_2]) = E_2 \cap ([E_1] \cap [E_2])$, i.e., if for all $e \in [E_1] \cap [E_2]$, $e \in E_1$ if and only if $e \in E_2$.

The *incomparability graph* $\tilde{G} = (V, \tilde{E})$ associated with a digraph (V, E) is defined such that for all $v, w \in V$,

$$(v, w) \in \tilde{E} \text{ if and only if neither } (v, w) \in E \text{ nor } (w, v) \in E.$$

Obviously, $[\tilde{E}] = \tilde{E}$.

A directed graph $G = (V, E)$ is said to be *transitive* if for all $x, y, z \in V$, $(x, y) \in E$ and $(y, z) \in E$ imply $(x, z) \in E$. Moreover, G is *acyclic* if for all $x_1, \ldots, x_k \in V$, $(x_1, x_2), (x_2, x_3), ..., (x_{k-1}, x_k) \in E$ implies $(x_k, x_1) \notin E$. Also G is *asymmetric* if $(v, w) \in E$ implies $(w, v) \notin E$. A *tournament* is an asymmetric graph (V, E) where E is complete, i.e., if for all distinct $v, w \in V$, either $(v, w) \in E$ or $(w, v) \in E$. Moreover, a digraph (V, E) is *transitively (re)orientable* if there exists a transitive and asymmetric graph (V, E') with $[E'] = [E]$. E' is also referred to as a *reorientation of* E.

The graphs in this paper are assumed to be induced by the preferences of a set of voters. Let $N = \{1, \ldots, n\}$ be a set of n voters and V a set of alternatives. The preferences of each voter i are given as *linear orders*, i.e., transitive, complete, and antisymmetric relations R_i over a set of alternatives V. A *preference profile* $R = (R_1, \ldots, R_n)$ associates a preference relation to each voter. Each preference profile gives rise to a *majority relation*, which holds between two alternatives v and w if the number of voters preferring v to w exceeds the number of voters preferring w to v. Formally define $n^R_{vw} = |\{i \in N: v\ R_i\ w\}|$, omitting the superscript when R is clear from the context. We say that (V, E) is *the majority graph* of preference profile R if

$$(v, w) \in E \text{ if and only if } n_{vw} > n_{wv}.$$

We also say that (V, E) *is induced by a k-voter profile* if (V, E) is the majority graph for some preference profile involving k voters. Defined thus, majority graphs are generally asymmetric. If the number of voters is odd, moreover, the majority graph is complete and therefore a tournament. We will also come to consider *weighted graphs* (V, w), where V is a set of vertices and $\mathrm{w}: V \times V \to \mathbb{Z}$ a weight function associating edge (v, w) with a weight. With a slight abuse of notation we also refer to weighted graphs as a pair (V, E), where the weight function is subsumed and it is understood that $E = \{(v, w): \mathrm{w}(v, w) > 0\}$. We say that a weighted graph (V, w) *is induced by* R if for all $v, w \in V$, $\mathrm{w}(v, w) = n_{vw} - n_{wv}$.

By a *voting rule* we understand a function that maps each preference profile to a subset of alternatives. Within the field of social choice theory a large number of voting rules have been proposed. The ones we will be concerned with in this paper are the *Banks set (BA)*, the *Tournament Equilibrium Set (TEQ)*, the *Slater set (SL)*, and *Ranked Pairs (RP)*.

3. CHARACTERIZATION RESULTS

Given an n-voter preference profile R, the *Pareto relation* holds between two alternatives v and w if all n voters prefer v to w. Dushnik and Miller [7] specified sufficient and necessary conditions for relations to be induced as the Pareto relation of a 2-voter preference profile. As for two voters the majority relation and the Pareto relation obviously coincide, we can rephrase their result for majority graphs as follows.

LEMMA 1. *A majority graph (V, E) is induced by a 2-voter preference profile if and only if it is transitive and its incomparability graph $(V, \tilde{E})$ is transitively orientable. Moreover, the weight of every edge is 2.*

If a graph (V, E) is induced by a 2-voter profile (R_1, R_2), then R_1 and R_2 coincide on E and are opposed on $\tilde{E}$, i.e., $R_1 \cap R_2 = E$. As R_1 and R_2 are both transitive, so is E. If E' is the respective reorientation of $\tilde{E}$, then $R_1 = E \cup E'$ and $R_2 = E \cup \overline{E'}$, or *vice versa.* Not all transitive graphs are induced by a 2-voter profile, as illustrated in Figure 1.

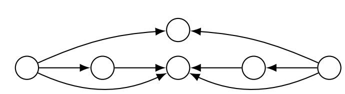

Figure 1: Example of a transitive graph that cannot be induced by a 2-voter profile

As a corollary of Lemma 1, we find that the union of pairwise disjoint graphs that are induced by 2-voter profiles is also induced by a 2-voter profile.

LEMMA 2. *Let $V_1, \ldots, V_k$ be pairwise disjoint and $(V_1, E_1), \ldots, (V_k, E_k)$ majority graphs induced by 2-voter profiles. Then, $(V_1 \cup \cdots \cup V_k, E_1 \cup \cdots \cup E_k)$ is also induced by a 2-voter profile.*

PROOF. Let $V = V_1 \cup \cdots \cup V_k$ and $E = E_1 \cup \cdots \cup E_k$ and consider the graph (V, E). As each of $(V_1, E_1), \ldots, (V_k, E_k)$ is induced by a 2-voter profile, by Lemma 1, each of $E_1, \ldots, E_k$ is transitive and each of $\tilde{E}_1, \ldots, \tilde{E}_k$ is transitively orientable. Let $E'_1, \ldots, E'_k$ be the respective transitive reorientations of $\tilde{E}_1, \ldots, \tilde{E}_k$. Since $V_1, \ldots, V_k$ are pairwise disjoint, $E_1 \cup \cdots \cup E_2$ can readily be seen to be transitive as well. Let furthermore $E^* = \bigcup_{1 \leq i < j \leq k} (V_i \times V_j)$. Observe that $\tilde{E} = [\tilde{E}_1] \cup \cdots \cup [\tilde{E}_k] \cup [E^*]$ and that $E'_1 \cup \cdots \cup E'_k \cup E^*$ is a transitive reorientation of $\tilde{E}$. The claim then follows by another application of Lemma 1. $\square$

Extensions of these results provide useful sufficient conditions for a graph to be induced by a constant larger number of voters. If the edge set of a graph can be decomposed into pairwise orientation compatible sets that satisfy the conditions of Lemma 1, the graph is induced by a profile with two voters per set.

LEMMA 3. *Let $(V, E_1), \dots, (V, E_k)$ be majority graphs induced by 2-voter profiles such that $E_1, \dots, E_k$ are pairwise orientation compatible. Then, $(V, E_1 \cup \dots \cup E_k)$ is induced by a $2k$-voter profile.*

PROOF. Let for each i with $1 \le i \le k$, (R_1^i, R_2^i) be a 2-voter profile that induces (V, E_i). By Lemma 1, for every $(v, w) \in E_i$ we know that both $v\ R_1^i\ w$ and $v\ R_2^i\ w$ and for every $(v, w) \notin E_i$, $v\ R_1^i\ w$ if and only if $w\ R_2^i\ v$. Now consider the preference profile $(R_1^1, R_2^1, \dots, R_1^k, R_2^k)$ and the majority graph (V, E) it induces. We argue that $E = E_1 \cup \dots \cup E_k$. First assume that $(v, w) \in E_i$ for some i with $1 \le i \le k$. Then, both $v\ R_1^i\ w$ and $v\ R_2^i\ w$. Since, $E_1, \dots, E_k$ are pairwise orientation compatible, $(w, v) \in E_j$ for no j with $1 \le j \le k$, i.e., for all j with $1 \le j \le k$ either $v\ R_1^j\ w$ and $v\ R_2^j\ w$ or $v\ R_1^j\ w$ if and only if $w\ R_2^j\ v$. It follows that a majority prefers v over w and thus $(v, w) \in E$. Now assume that $(v, w) \in E_i$ for no i with $1 \le i \le k$. Then for all i with $1 \le i \le k$ either both $w\ R_1^i\ v$ and $w\ R_2^i\ v$ or $w\ R_1^j\ v$ if and only if $v\ R_2^j\ w$. It is seen easy to see that v is not majority preferred to w, i.e., $(v, w) \notin E$. $\square$

In the above proof, for the majority graph (V, E_i) induced by (R_1^i, R_2^i), we have for every $v, w \in V$,

$$n_{vw} - n_{wv} = \begin{cases} 2 & \text{if } (v,w) \in E_i, \\ -2 & \text{if } (w,v) \in E_i, \text{ and} \\ 0 & \text{if } (v,w) \in \tilde{E}_i. \end{cases}$$

Observing that $E_1, \dots, E_k$ are pairwise orientation compatible, it can then be appreciated that for the majority graph (V, E) induced by the profile $(R_1^1, R_2^1, \dots, R_1^k, R_2^k)$ we have for every $v, w \in V$,

$$n_{vw} - n_{wv} = \begin{cases} 2 \cdot |\{E_i : (v,w) \in E_i\}| & \text{if } (v,w) \in E, \\ -2 \cdot |\{E_i : (v,w) \in E_i\}| & \text{if } (w,v) \in E, \text{ and} \\ 0 & \text{if } (v,w) \in \tilde{E}_i. \end{cases}$$

Apart from an example of a tournament of order eight that cannot be obtained using three voters [16], little was known about the majority graphs that are induced by 3-voter profiles. In a much similar vein as Lemma 1, we now provide a characterization of these graphs.

LEMMA 4. *A tournament (V, E) is induced by a 3-voter profile if and only if there are disjoint sets E_1, E_2 with $E = E_1 \cup E_2$ such that E_1 is transitive and E_2 is both acyclic and transitively reorientable. Then, the weight of every edge in E_1 is either 1 or 3 and that of each edge in E_2 is 1.*

PROOF. For the if-direction, assume that there are disjoint sets E_1, E_2 with $E = E_1 \cup E_2$ such that E_1 is transitive and E_2 is both acyclic and transitively reorientable. Consider the graph (V, E_1) and observe that for the corresponding incomparability graph $(V, \tilde{E}_1)$, $\tilde{E}_1 = [E_2]$. It follows that $\tilde{E}_1$ is transitively orientable and, by Lemma 1, that (V, E_1) is induced by a 2-voter profile (R_1, R_2) and that all edges in E_1 have weight 2. As E_2 is acyclic, there is a (strict) preference relation R_3 with $E_2 \subseteq R_3$. Now consider the majority graph induced by the preference profile (R_1, R_2, R_3), which apparently coincides with (V, E). E_1 is determined by R_1 and R_2 and each of its edges obtains weight 1 or 3 depending on whether R_3 agrees with both R_1 and R_2 or not. Moreover, E_2 is determined by R_3, as R_1 and R_2 can be assumed to specify contrary preferences on this part.

For the only-if-direction, assume that (V, E) is the majority graph induced by the 3-voter preference profile (R_1, R_2, R_3). Let furthermore (V, E_1) be the majority graph induced by (R_1, R_2) and $E_2 = R_3 \cap ((V \times V) \setminus [E_1])$. By Lemma 1, (V, E_1) is transitive and $\tilde{E}_1$ is transitively (re)orientable, where $(V, \tilde{E}_1)$ is the incomparability graph of (V, E_1). As R_3 is transitive (and strict) E_2 is obviously acyclic. Observe furthermore that $[R_3 \cap ((V \times V) \setminus [E_1])] = [\tilde{E}_1]$. It follows that E_2 is transitively reorientable. $\square$

The if-direction of Lemma 4 can also obtained as a special case of the following lemma, which provides sufficient conditions for a graph to be induced by profiles involving an odd number of voters.

LEMMA 5. *Let (V, E) be a tournament and $(V, E_1), \dots, (V, E_k)$ be majority graphs induced by 2-voter profiles such that $E_1, \dots, E_k$ are orientation compatible. Let, moreover, $E_{k+1} = E \setminus (E_1 \cup \dots \cup E_k)$ be acyclic. Then, (V, E) is induced by a $2k + 1$-voter profile.*

PROOF. In virtue of Lemma 3 we know that $(V, E_1 \cup \dots \cup E_k)$ is induced by a $2k$-voter profile $(R_1, \dots, R_{2k})$. Inspection of the proof also reveals that every edge $(v, w) \in E_1 \cup \dots \cup E_k$ has a positive even weight. As E_{k+1} is acyclic and asymmetric, there is some (strict) preference relation R_{2k+1} with $E_{k+1} \subseteq R_{2k+1}$. It can then easily be seen that the majority graph induced by $(R_1, \dots, R_{2k}, R_{2k+1})$ equals (V, E), $E \setminus E_{k+1}$ being determined by majorities of at least two in $(R_1, \dots, R_{2k})$ and E_{k+1} by R_{2k+1}, each edge in which has then weight 1. $\square$

4. THE BANKS SET

The *Banks set*, a concept proposed by Jeffrey Banks, associates with each majority tournament the maximal elements of its maximal (with respect to set-inclusion) transitive subtournaments (see, e.g., [13]).

Although finding a random alternative in the Banks set can be achieved in polynomial time, deciding whether an alternative belongs to the Banks set is NP-complete as shown by Woeginger [19]. Brandt et al. [3] gave an arguably simpler proof of this result by a reduction from 3SAT: every formula φ in 3CNF can be transformed in polynomial time into a tournament T_φ^{BA} with a decision node c_0 such that c_0 is in the Banks set of T_φ^{BA} if and only if φ is satisfiable. We may assume that no propositional variable occurs, negatively or positively, more than once in each clause. We have P denote the set of variables of the propositional language in which φ is formulated.

Let $\mathcal{G}^{BA}$ denote the class of tournaments thus constructed for formulas in 3CNF. We show that every tournament in this class $\mathcal{G}^{BA}$ is induced by a 7-voter profile, proving that deciding whether an alternative is in the Banks set is already hard if there are only seven voters.

A tournament (V, E) is in *the class* $\mathcal{G}^{BA}$ if it satisfies the following properties. There is an odd integer m such that,

$$V = C \cup U_1 \cup \dots \cup U_m,$$

where $C, U_1, \ldots, U_m$ are pairwise disjoint and $C = \{c_0, \ldots, c_m\}$. We have C_i denote the singleton $\{c_i\}$. If i is odd, $U_i = \{u_i^1, u_i^2, u_i^3\}$ whereas if i is even U_i is a singleton $\{u_i\}$. Let $X = \bigcup\{U_i : i \text{ is odd}\}$ and $Y = \bigcup\{U_i : i \text{ is even}\}$.

Intuitively, (V, E) is T_φ^{BA} for some φ in 3CNF with $\frac{1}{2}(m+1)$ clauses. If i is odd, U_i corresponds to a clause of φ and the nodes it contains represent (tokens of) literals. We assume each of these nodes u_i^j to be labeled by the literal $\lambda(u_i^j)$ it represents. For *odd* $i \in \{1, \ldots, m\}$ and $j \in \{1, 2, 3\}$ we define,

$$\begin{aligned} U_i^j &= \{u_i^j\} \\ U_i^p &= \{u \in U_i : \lambda(u) = p\} \\ U_i^{\neg p} &= \{u \in U_i : \lambda(u) = \neg p\} \end{aligned}$$

Moreover, for *even* $i \in \{1, \ldots, m\}$ and $j \in \{1, 2, 3\}$, we stipulate,

$$U_i^j = U_i^p = U_i^{\neg p} = \emptyset.$$

Observe that $\bigcup_{\substack{p \in P \\ 1 \le i \le m}} (U_i^p \cup U_i^{\neg p}) = X$.

We are now in a position to define the edge set E.

$$\begin{aligned} E = &\bigcup_{i<j} (C_j \times C_i) \cup \bigcup_{i \ne j} (C_i \times U_j) \cup \bigcup_{i=j} (U_j \times C_i) \cup \\ &\bigcup_{1 \le i \le m} \big((U_i^1 \times U_i^2) \cup (U_i^2 \times U_i^3) \cup (U_i^3 \times U_i^1)\big) \cup \\ &\bigcup_{i<j} \big((U_i \times U_j) \setminus \overline{E^\varphi}\big) \cup E^\varphi, \end{aligned}$$

where

$$E^\varphi = \bigcup_{\substack{p \in P \\ i>j}} (U_i^p \times U_j^{\neg p}) \cup \bigcup_{\substack{p \in P \\ i>j}} (U_i^{\neg p} \times U_j^p).$$

Figure 2 illustrates this type of tournament. We also refer to E^φ as the *formula dependent* or the *flesh* of the tournament T_φ^{BA}. The edge set

$$(E \setminus E^\varphi) \cup \overline{E^\varphi}$$

we refer to as its *skeleton*. In the theorem that follows we show that the skeleton of each tournament T_φ^{BA} is induced by a 3-voter profile such that the edges in $\overline{E^\varphi}$ all get weight 1, whereas E^φ can be partitioned in two (orientation compatible) edge sets, $\bigcup_{\substack{p \in P \\ i>j}} (U_i^p \times U_j^{\neg p})$ and $\bigcup_{\substack{p \in P \\ i>j}} (U_i^{\neg p} \times U_j^p)$, both of which are induced by a 2-voter profile and all edges get weight 2. Some reasoning and an application of Lemma 5 then gives the desired result.

THEOREM 1. *Computing the Banks set is NP-hard, if the number of voters is at least seven.*

PROOF. Let (V, E) be a tournament in $\mathcal{G}^{BA}$. It suffices to show that (V, E) is induced by a 7-voter profile. To this end define:

$$\begin{aligned} E_1 &= \bigcup_{i>j} (C_i \times C_j) \cup \bigcup_{i\,>\,j} (C_i \times U_j) \cup \bigcup_{1 \le i \le m} (U_i^3 \times U_i^1), \\ E_2 &= \bigcup_{\substack{p \in P \\ i>j}} (U_i^p \times U_j^{\neg p}), \\ E_3 &= \bigcup_{\substack{p \in P \\ i>j}} (U_i^{\neg p} \times U_j^p), \text{ and} \\ E_4 &= E \setminus (E_1 \cup E_2 \cup E_3). \end{aligned}$$

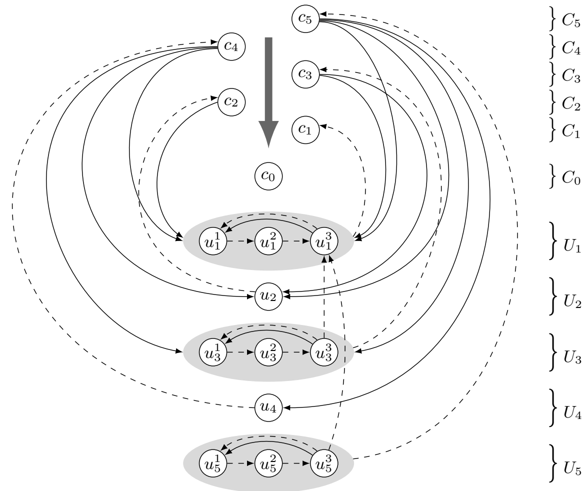

Figure 2: A tournament $T_\varphi^{BA} = (V, E)$ in the class $\mathcal{G}^{BA}$, where E is given by dashed edges and it is understood that missing edges point downwards. Moreover, $\lambda(u_5^3) = \lambda(u_3^3) = \overline{\lambda}(u_1^3)$. The solid edges, including all edges (c_i, c_j) for $i > j$ represented by the gray arrow, correspond to the edge set E_1 in Theorem 1.

(For E_1, also see Figure 2.) Observe that $E = E_1 \cup E_2 \cup E_3 \cup E_4$ and that E_1, E_2, and E_3 are pairwise orientation compatible. In virtue of Lemma 5, it therefore suffices to show that (V, E_1), (V, E_2), and (V, E_3) are induced by 2-voter profiles and that (V, E_4) is acyclic.

By making the obvious case distinctions, it is easy but tedious to show that (V, E_1) is transitive. Now let,

$$\begin{aligned} E_1' = &\bigcup_{i \le j} (C_i \times U_j) \cup \bigcup_{i<j} (U_i \times U_j) \cup \\ &\bigcup_{1 \le i \le m} \big((U_i^2 \times U_i^1) \cup (U_i^2 \times U_i^3)\big). \end{aligned}$$

It can readily be appreciated that E_1' is a transitive orientation of $\tilde{E}_1$. Hence, by Lemma 1, (V, E_1) is induced by a 2-voter profile.

To see that (V, E_2) is induced by a 2-voter profile as well, define for each propositional variable p,

$$E_2^{p,\neg p} = \bigcup_{i>j} (U_i^p \times U_j^{\neg p}).$$

Consider the subgraphs $(X, E_2^{p,\neg p})$ for the propositional variables p in P along with $(C \cup Y, \emptyset)$. Obviously, the vertex sets of these subgraphs are pairwise disjoint. Also observe that $V = C \cup Y \cup X$ and $E_2 = \bigcup_{p \in P} E_2^{p,\neg p}$. Moreover, for each p, $E_2^{p,\neg p}$ is obviously transitive as there are no $v, w, u \in V$ with both $(v, w), (w, u) \in E_2^{p,\neg p}$. Furthermore, one can easily check that

$$\bigcup_{i<j} \big((U_i^p \times U_j^{\neg p}) \cup (U_i^p \times U_j^p) \cup (U_i^{\neg p} \times U_j^{\neg p})\big)$$

is a reorientation of $\tilde{E}_2^{p,\neg p}$ (recall that the literals with which the nodes in the same layer are labeled all involve different propositional variables). As $(C \cup Y, \emptyset)$ is obviously transitive and every transitive relation over $C \cup Y$ is a reorientation

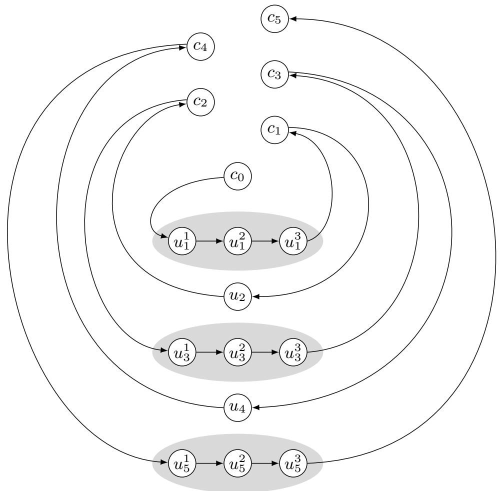

Figure 3: E_4 is contained in the transitive closure E_4' of the ordering depicted. Therefore, E_4 is acyclic.

of $\tilde{\emptyset}$, it follows from Lemmas 1 and 3 that (V, E_2) is induced by a 2-voter profile. The argument that the same holds for (V, E_3) runs along analogous lines.

Finally, to see that E_4 is acyclic, observe that E_4 can equivalently be written as,

$$E_4 = \bigcup_{i<j} \big(((U_i \times U_j) \setminus (\overline{E_2} \cup \overline{E_3})) \cup (C_i \times U_j)\big) \cup \bigcup_{1\le i\le m} (U_i \times C_i) \cup \bigcup_{1\le i\le m} \big((U_i^1 \times U_i^2) \cup (U_i^2 \times U_i^3)\big).$$

Now define E_4' as follows.

$$E_4' = \bigcup_{i<j} \big((U_i \times U_j) \cup (C_i \times U_j) \cup (C_i \times C_j)\big) \cup \bigcup_{i\le j} (U_i \times C_j) \cup \bigcup_{1\le i\le m} \big((U_i^1 \times U_i^2) \cup (U_i^2 \times U_i^3) \cup (U_i^1 \times U_i^3)\big).$$

Some reflection reveals that (V, E_4') is a transitive tournament, which defines an ordering over V, whose initial part is like

$$c_0, u_1^1, u_1^2, u_1^3, c_1, u_2, c_2, u_3^1, u_3^2, u_3^3, c_3, u_4, c_4, \ldots.$$

It can easily be seen that $E_4 \subseteq E_4'$ and, thus, that (V, E_4) is acyclic. $\square$

5. THE TOURNAMENT EQUILIBRIUM SET

The *tournament equilibrium set* (*TEQ*) is another function that, like the Banks set, selects a set of alternatives from each tournament (see, e.g., [13]). Its recursive definition is based on the notion of retentiveness. Given a tournament (V, E), a subset $X \subseteq V$ is said to be *TEQ-retentive* if for all $v \in X$ all alternatives chosen by *TEQ* from the subtournament of (V, E) induced by $\{w \in V : (w, v) \in E\}$ are contained in X. *TEQ* is then defined so as to select the union of the inclusion-minimal *TEQ*-retentive sets from each tournament.

Brandt et al. [3] have shown that computing *TEQ* is NP-hard by a reduction from 3SAT. For every formula φ in 3CNF a tournament T_φ^{TEQ} can be constructed such that *TEQ* selects a decision node c_0 from T_φ^{TEQ} if and only if φ is satisfiable. The class of these tournaments T_φ^{TEQ} is denoted by $\mathcal{G}^{TEQ}$ and the tournaments in this class bear a strong structural similarity to those in $\mathcal{G}^{BA}$, which can be exploited to show that, similarly as for the Banks set, every tournament in $\mathcal{G}^{TEQ}$ is induced by a 7-voter profile. It follows that computing *TEQ* is already hard if the number of voters is seven. Apart from a number of tedious details, the proof of this result runs along analogous lines as that of Theorem 1 for the Bank set and is therefore omitted.

THEOREM 2. *Computing* TEQ *is NP-hard, if the number of voters is at least seven.*

6. THE SLATER SET

Slater's rule considers rankings over alternatives, which minimally conflict with the pairwise majority relation. The set of maximal elements of these rankings is known as the Slater set (see, e.g., [13]). The close relationship between Slater rankings and feedback arc sets can be used to easily show that computing Slater rankings is NP-hard in general digraphs. It was proved by Alon [1] and Conitzer [6] that computing feedback arc sets is NP-hard even in tournaments. As a consequence, computing Slater rankings and the Slater set is NP-hard as well [12]. We will analyze the proof of Conitzer [6], a reduction from MAXSAT. The latter problem asks for an assignment to the propositional variables in a Boolean formula φ such that at least a given number s_1 of clauses is satisfied. In the reduction, a corresponding tournament T_φ^{SL} is constructed for which a Slater ranking with at most s_2 inconsistent edges exists if and only if such an assignment for φ exists, where s_2 depends on φ and s_1. Let $\mathcal{G}^{SL}$ denote the class of all tournaments T_φ^{SL} obtained from a Boolean formula φ according to this construction. We will show that every tournament in $\mathcal{G}^{SL}$ is induced by a 9-voter profile, proving that finding a Slater ranking is already hard if there are only nine voters.

A tournament (V, E) is in the class $\mathcal{G}^{SL}$ if it satisfies the following properties. There exist integers $m, l \geq 1$, such that

$$V = C \cup \bigcup_{\substack{1\le i\le m\\ 1\le j\le 6}} T_i^j,$$

where C and all T_i^j are pairwise disjoint and for $1 \le i \le m$

$$C = \{c_1, \ldots, c_{|C|}\},$$
$$T_i^j = \{t_i^{j,1}, \ldots, t_i^{j,l}\}.$$

Every $T_i^1, \ldots, T_i^6$ correspond to a variable in φ whose clauses are represented by the vertices in C. Each subtournament $\big(T_i^j, E \cap \big(T_i^j \times T_i^j\big)\big)$ has to be a transitive component, i.e., it is a linear order and for a vertex $v \in V \setminus T_i^j$ and vertices $v_1, v_2 \in T_i^j$, either $\{(v_1, v), (v_2, v)\}$ or $\{(v, v_1), (v, v_2)\}$ have to be in E. Therefore, we can treat T_i^j as a single vertex denoted by t_i^j. For notational convenience, we also write T^j for $\bigcup_{1\le i\le m} t_i^j$. For (V, E) to be in $\mathcal{G}^{SL}$, the edge set has to be of the form $E = E^\sigma \cup E^\varphi$, where E^σ (the *skeleton*) and E^φ (the *formula dependent part*) are disjoint. The skeleton

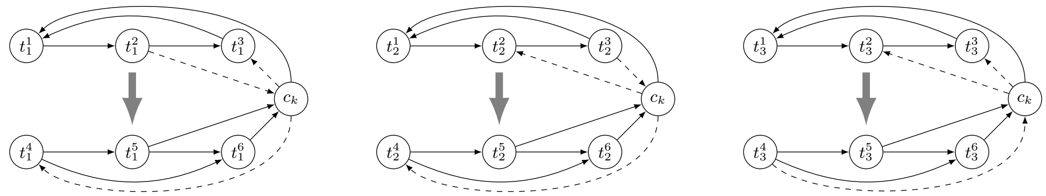

Figure 4: Illustration of the hardness construction for computing a Slater ranking: three cases for the formula-dependent dashed edges between a variable gadget on nodes $t_i^1, \dots, t_i^6$ and a clause node c_k. The omitted edges inside a gadget are not formula-dependent and are meant to point downwards as indicated by the solid gray arrows.

is again composed as

$$E^\sigma = E_A^\sigma \cup E_B^\sigma \cup E_C^\sigma \cup E_D^\sigma \cup E_E^\sigma \cup E_F^\sigma$$

such that

$$\begin{aligned}
E_A^\sigma &= \{(c_{k_1}, c_{k_2}) \in C \times C \colon k_1 < k_2\},\\
E_B^\sigma &= \bigcup_{1 \le i \le m} \{(t_i^1, t_i^2), (t_i^2, t_i^3), (t_i^3, t_i^1)\}\\
E_C^\sigma &= \bigcup_{1 \le i \le m} \{(t_i^4, t_i^5), (t_i^5, t_i^6), (t_i^4, t_i^6)\}\\
E_D^\sigma &= \bigcup_{\substack{1 \le i \le m\\ j \in \{1,2,3\}\\ j' \in \{4,5,6\}}} \{(t_i^j, t_i^{j'})\} \cup \bigcup_{\substack{1 \le i, i' \le m\\ 1 \le j, j' \le 6}} \{(t_i^j, t_{i'}^{j'}) \colon i < i'\}\\
E_E^\sigma &= (C \times T^1) \cup (T^5 \times C) \cup (T^6 \times C), \text{ and}\\
E_F^\sigma &= \bigcup_{\substack{1 \le i \le m\\ 1 \le j \le 6}} \{(t_i^{j,l_1}, t_i^{j,l_2}) \colon l_1 < l_2\}.
\end{aligned}$$

The formula dependent part E^φ specifies all edges between $\{T^2, T^3, T^4\}$ and C. The last condition is that for each pair (i, k) exactly one of the three possible edges $(t_i^2, c_k), (t_i^3, c_k), (t_i^4, c_k)$ is in E with the two other pointing in the other direction. The different cases are shown in Figure 4. Note that (V, E) is indeed a tournament.

THEOREM 3. *Computing the Slater set is NP-hard, if the number of voters is at least nine.*

PROOF. Let (V, E) be a tournament in $\mathcal{G}^{SL}$. As in the proof of Theorem 1, we decompose the edge set into disjoint sets E_1, E_2, E_3, E_4, E_5 and claim that they satisfy the prerequisites of Lemma 5. Let

$$\begin{aligned}
E_1 &= \{(t_i^1, t_i^2) \in E \colon 1 \le i \le m\} \cup E_C^\sigma \cup E_D^\sigma,\\
E_2 &= \big((C \times T^2) \cup (C \times T^4)\big) \cap E,\\
E_3 &= \big((C \times T^3) \cup (C \times T^4)\big) \cap E,\\
E_4 &= \big((C \times T^2) \cup (C \times T^3)\big) \cap E, \text{ and}\\
E_5 &= E \setminus (E_1 \cup E_2 \cup E_3 \cup E_4).
\end{aligned}$$

We claim that (V, E_5) is acyclic and that $(V, E_1), (V, E_2), (V, E_3)$, and (V, E_4) are induced by 2-voter profiles and orientation compatible. Due to space restrictions we omit the proofs. It follows from Lemma 5 that (V, E) is induced by a 9-voter profile.

The fact that computing the Slater set is NP-hard for any even number of voters greater than two follows from results by Dwork et al. [8] and Biedl et al. [2]. □

7. RANKED PAIRS

The last voting rule investigate is the *ranked pairs method (RP)*. There are two versions of RP commonly discussed in the literature. The one we are concerned about is the *neutral* one, i.e., the one that does not discriminate among the alternatives. Brill and Fischer [5] have recently shown that deciding whether a given alternative is a winner according to this version of RP is NP-complete.

RP is usually regarded as a procedure. First, it defines a priority over all pairs of alternatives, and then ranks the alternatives iteratively in order of priority. The priority over pairs (a, b) of alternatives is defined based on the number of voters who prefer a to b. To avoid creating cycles, any pair whose addition would yield a cycle is discarded in the procedure. The neutral version of RP, which was defined by Tideman [18] and considered by Brill and Fischer [5], returns the set of all rankings the above procedure gives for *some* tie breaking rule. From this point on, we refer to this variant simply by RP.

The NP-hardness proof by Brill and Fischer [5] is by a reduction from the Boolean satisfiability problem (SAT). For each Boolean formula φ in CNF they constructed a weighted graph G_φ^{RP} such that a decision node d is selected by RP from G_φ^{RP} if and only if φ is satisfiable. The construction, of course, works just as well for a reduction from 3SAT. We may also assume that in every formula φ in 3CNF no variable occurs more than once in each clause.

We first define the class $\mathcal{G}^{RP}$ in which the weighted graphs G_φ^{RP} for formulas φ in 3CNF are contained. Later we prove that every graph in this class is induced by an 8-voter profile, showing that deciding whether a given alternative is a ranked pairs winner is NP-complete.

A weighted graph (V, E) (with weight function w) belongs to $\mathcal{G}^{RP}$ if and only if it fulfills the following conditions. There are some integers $k, m \ge 1$ such that

$$V = D \cup U_1 \cup \cdots \cup U_m \cup X_1 \cup \cdots \cup X_k,$$

where, for $1 \le i \le m$ and $1 \le j \le k$,

$$\begin{aligned}
D &= \{d\},\\
U_i &= \{u_i^1, u_i^2, u_i^3, u_i^4\}, \text{ and}\\
X_j &= \{x_j\}.
\end{aligned}$$

If (V, E) is obtained as the graph G_φ^{RP} for some φ in 3CNF, k is the number of clauses, m the number of variables occurring in φ, the U_is are the variable gadgets, the X_js the clause gadgets, and, finally, D the decision node. Let $U^j = \bigcup_{i=1}^m \{u_i^j\}$, $U = \bigcup_{i=1}^m U_i$ and $X = \bigcup_{i=1}^k X_i$. Moreover, $E = E^\sigma \cup E^\varphi$, where E^σ (the *skeleton*) and E^φ (the *formula dependent*

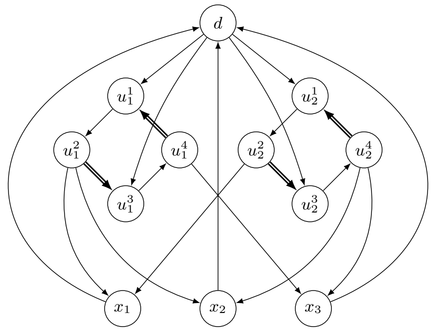

Figure 5: A graph (V, E) in the class $\mathcal{G}^{RP}$. The thick edges have weight 4 whereas the thin edges have weight 2.

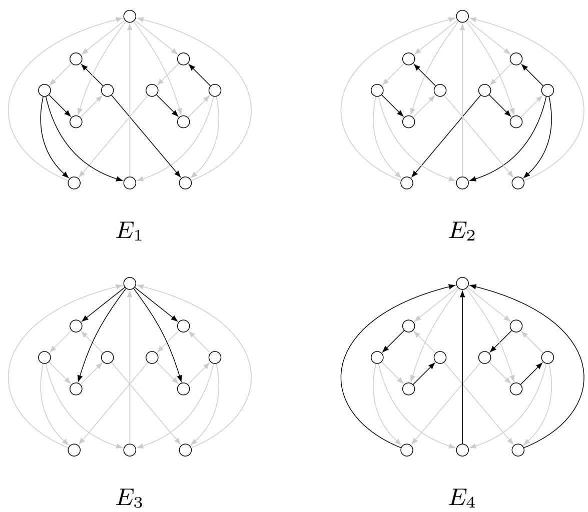

Figure 6: The sets E_1, E_2, E_3, and E_4 for the graph of Figure 5 as defined in the proof of Theorem 4.

part) are disjoint such that

$$E^\sigma = (D \times (U^1 \cup U^3)) \cup (X \times D) \cup \bigcup_{i=1}^{m} \{(u_i^1, u_i^2), (u_i^2, u_i^3), (u_i^3, u_i^4), (u_i^4, u_i^1)\}$$

and E^φ is such that for all $1 \leq i \leq m$ and all $1 \leq j \leq k$:

$$E^\varphi \subset (U^2 \cup U^4) \times X,$$
$$|E^\varphi \cap (U^2 \cup U^4) \times X_j)| \leq 3, \text{ and}$$
$$|E^\varphi \cap (U_i^2 \cup U_i^4) \times X_j)| \leq 1,$$

i.e., every vertex in X has at most three incoming edges (intuitively corresponding to the literals x contains) and at most one from every U_i (intuitively corresponding to that no propositional variable occurs more than once in each clause). Finally, the weight function w is defined such that all edges in $E \cap ((U^2 \times U^3) \cup (U^4 \times U^1))$ have weight 4 and all edges in $E \setminus ((U^2 \times U^3) \cup (U^4 \times U^1))$ have weight 2. The reader is deferred to Figure 5 for an example illustrating this definition of the class $\mathcal{G}^{RP}$.

Not being a complete graph, G^{RP} can only be induced by a profile involving an even number of voters. Actually we will prove that only eight voters suffice to induce any graph in $\mathcal{G}^{RP}$. In order to do so we rely on the graph-theoretical concept of a forest of stars. A graph (V, E) is a *(directed) star* if there is some $v^* \in V$ such that $E = \{(v^*, v) \colon v \in V \setminus \{v^*\}\}$ or $\overline{E} = \{(v^*, v) \colon v \in V \setminus \{v^*\}\}$. Vertex v^* is also called the *center* and the other vertices *leaves*. (V, E) is said to be a *forest of (directed) stars* if there is a partitioning $\{V_1, \ldots, V_k\}$ of V and a partitioning $\{E_1, \ldots, E_k\}$ of E such that each (V_i, E_i) is a star.

Lemma 6. *Every forest of stars is induced by a 2-voter profile.*

We are now in a position to prove the main result of this section.

Theorem 4. *Deciding whether a given alternative is a ranked pairs winner is NP-complete, if the number of voters is even and at least eight.*

Proof. Let (V, E) be a graph (with weight function w) in $\mathcal{G}^{RP}$. Intuitively, $(V, E) = G_\varphi^{RP}$ for some formula φ in 3CNF. It suffices to show that (V, E) is induced by an 8-voter profile. As an auxiliary notion, let for each $1 \leq j \leq k$,

$$E^\varphi \cap ((U^2 \cup U^4) \times X_j) = E_{j,1}^\varphi \cup E_{j,2}^\varphi \cup E_{j,3}^\varphi,$$

where $|E_{j,i}^\varphi| \leq 1$ for all $1 \leq i \leq 3$. Intuitively, $E_{j,1}^\varphi$, $E_{j,2}^\varphi$, and $E_{j,3}^\varphi$ impose an ordering on the incoming edges of vertex x_j. Also set

$$E_i^\varphi = \bigcup_{j=1}^{k} E_{j,i}^\varphi$$

for each $1 \leq i \leq 3$, i.e., E_i^φ collects the i-th incoming edges of the vertices in X. Now define the following edge sets.

$$E_1 = E_1^\varphi \cup \bigcup_{i=1}^{m} \{(u_i^2, u_i^3), (u_i^4, u_i^1)\},$$
$$E_2 = E_2^\varphi \cup \bigcup_{i=1}^{m} \{(u_i^2, u_i^3), (u_i^4, u_i^1)\},$$
$$E_3 = E_3^\varphi \cup (D \times (U^1 \cup U^3)), \text{ and}$$
$$E_4 = (X \times D) \cup \bigcup_{i=1}^{m} \{(u_i^1, u_i^2), (u_i^3, u_i^4)\}.$$

Observe that $E = E_1 \cup E_2 \cup E_3 \cup E_4$. Moreover, each of (V, E_1), (V, E_2), (V, E_3), and (V, E_4) is a forest of trees. Hence, by Lemma 6 we may assume they are induced by the 2-voter profiles (R_1^1, R_2^1), (R_1^2, R_2^2), (R_1^3, R_2^3), and (R_1^4, R_2^4), respectively. Moreover, E_1, E_2, E_3, and E_4 all contained in E and therefore also pairwise orientation compatible. By inspection of the proof of Lemma 3 it thus follows that (V, E) is induced by the 8-voter profile

$$R = (R_1^1, R_2^1, R_1^2, R_2^2, R_1^3, R_2^3, R_1^4, R_2^4).$$

Moreover, E_1, E_3, and E_4 as well as E_1, E_3, and E_4 are pairwise disjoint whereas $E_1 \cap E_2 = \bigcup_{i=1}^{m} \{(u_i^2, u_i^3), (u_i^4, u_i^1)\}$. Thus, by the remark following Lemma 3, all edges in $E \setminus \bigcup_{i=1}^{m} \{(u_i^2, u_i^3), (u_i^4, u_i^1)\}$ have weight 2, whereas those in $\bigcup_{i=1}^{m} \{(u_i^2, u_i^3), (u_i^4, u_i^1)\}$ have weight 4. We may conclude that also the graph (V, E) with its weights is induced by the 8-voter profile R. $\square$

Voting rule	NP-hard for $n \geq$
Banks set	7 voters
Tournament equilibrium set	7 voters
Slater set	9 voters
Ranked pairs	8 voters

Table 1: Numbers of voters for which winner determination is NP-hard. The Banks set and the tournament equilibrium set are defined for an odd number of voters only. Computing the Slater set for an even number of voters is NP-hard when $n \geq 4$ [8, 2]. Our result for ranked pairs is for an even number of voters.

8. CONCLUSION

Many hardness results in computational social choice only hold if the number of voters is roughly of the same order as the number of alternatives. In many applications, however, the number of voters can be much smaller than the number of alternatives and it is unclear whether hardness still holds.

We gave complete characterizations of majority graphs induced by two and three voters, respectively, and provided sufficient conditions for majority graphs to be induced by k voters. We then leveraged these conditions to show that winner determination for the Banks set, the tournament equilibrium set, Slater's rule, and ranked pairs remains hard even when there is only a small constant number of voters. This was achieved by analyzing existing hardness proofs and checking whether the class of majority graphs used in these constructions can be induced by small constant numbers of voters. Our results are summarized in Table 1.

We believe there is some very interesting potential for future work. First, and most importantly, it would be desirable to completely characterize the sets of graphs inducible by four, five, or more voters. Furthermore, it would be interesting to investigate whether it can be checked in polynomial time whether a graph is induced by a given number of voters. It follows from results by Pnueli et al. [15] and Yannakakis [20] that this is possible for graphs induced by 2-voter profiles. For three, however, this problem is already open. Finally, our techniques can be applied to verify whether other existing hardness proofs in computational social choice remain intact for a bounded number of voters. This would be particularly interesting for Kemeny's rule (which is only known to be hard for a constant *even* number of voters greater than two) and hardness shields against manipulation, bribery, and control.

9. ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft under grants BR 2312/7-1 and BR 2312/9-1. Paul Harrenstein was also partly supported by the ERC under Advanced Grant 291528 ("RACE").

10. REFERENCES

[1] N. Alon. Ranking tournaments. *SIAM Journal on Discrete Mathematics*, 20(1):137–142, 2006.

[2] T. Biedl, F. J. Brandenburg, and X. Deng. On the complexity of crossings in permutations. *Discrete Mathematics*, 309(7):1813–1823, 2009.

[3] F. Brandt, F. Fischer, P. Harrenstein, and M. Mair. A computational analysis of the tournament equilibrium set. *Social Choice and Welfare*, 34(4):597–609, 2010.

[4] F. Brandt, V. Conitzer, and U. Endriss. Computational social choice. In G. Weiß, editor, *Multiagent Systems*. MIT Press, 2nd edition, 2013. Forthcoming.

[5] M. Brill and F. Fischer. The price of neutrality for the ranked pairs method. In J. Hoffmann and B. Selman, editors, *Proc. of 26th AAAI Conference*, pages 1299–1305. AAAI Press, 2012.

[6] V. Conitzer. Computing Slater rankings using similarities among candidates. In *Proc. of 21st AAAI Conference*, pages 613–619. AAAI Press, 2006.

[7] B. Dushnik and E. W. Miller. Partially ordered sets. *American Journal of Mathematics*, 63(3):600–610, 1941.

[8] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank aggregation methods for the web. In *Proceedings of the 10th International Conference on the World Wide Web (WWW)*, pages 613–622. ACM Press, 2001.

[9] P. Erdős and L. Moser. On the representation of directed graphs as unions of orderings. *Publ. Math. Inst. Hungar. Acad. Sci.*, 9:125–132, 1964.

[10] P. Faliszewski, E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. A richer understanding of the complexity of election systems. In S. Ravi and S. Shukla, editors, *Fundamental Problems in Computing: Essays in Honor of Professor Daniel J. Rosenkrantz*. Springer, 2009.

[11] O. Hudry. NP-hardness results for the aggregation of linear orders into median orders. *Annals of Operations Research*, 163(1):63–88, 2008.

[12] O. Hudry. On the complexity of Slater's problems. *European Journal of Operational Research*, 203(1): 216–221, 2010.

[13] J.-F. Laslier. *Tournament Solutions and Majority Voting*. Springer, 1997.

[14] D. C. McGarvey. A theorem on the construction of voting paradoxes. *Econometrica*, 21(4):608–610, 1953.

[15] A. Pnueli, A. Lempel, and S. Even. Transitive orientation of graphs and identification of permutation graphs. *Canadian Journal of Mathematics*, 23: 160–175, 1971.

[16] D. Shepardson and C. A. Tovey. Smallest tournament not realizable by $\frac{2}{3}$-majority voting. *Social Choice and Welfare*, 33(3):495–503, 2009.

[17] R. Stearns. The voting problem. *American Mathematical Monthly*, 66(9):761–763, 1959.

[18] T. N. Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3): 185–206, 1987.

[19] G. J. Woeginger. Banks winners in tournaments are difficult to recognize. *Social Choice and Welfare*, 20: 523–528, 2003.

[20] M. Yannakakis. The complexity of the partial order dimension problem. *SIAM Journal on Algebraic and Discrete Methods*, 3(3):351–358, 1982.