Sequential majority voting with incomplete profiles

M. S. Pini, F. Rossi and **K. B. Venable**
University of Padova, Italy
email: {mpini,frossi,kvenable}@math.unipd.it

T. Walsh
NICTA and UNSW Sydney, Australia
email: Toby.Walsh@nicta.com.au

Abstract

In sequential majority voting, preferences are aggregated by a sequence of pairwise comparisons (also called an agenda) between candidates. The result of each comparison is determined by a weighted majority vote between the agents. In this paper we consider the situation where the agents may not have revealed all their preferences. This is common in real-life settings, due to privacy issues or an ongoing elicitation process. We study the computational complexity of determining the winner(s), given that some preferences may not yet be revealed and the agenda is not yet known or decided. We show that it is easy to determine if a candidate must win whatever the agenda. On the other hand, it is hard to know whether a candidate can win in at least one agenda for at least one completion of unrevealed preferences. This is also true if the agenda is balanced (that is, each candidate must win the same number of pairwise competitions to win overall).

Introduction

A general method for aggregating preferences in a multi-agent systems is running an election between the different options using a voting rule. Unfortunately, eliciting preferences from agents to be able to run such an election is a difficult, time-consuming and costly process. Agents may also be unwilling to reveal all their preferences for privacy reasons. Fortunately, we can often determine the outcome a long time before all the preferences have been revealed (Conitzer & Sandholm 2002b). For example, it may be that one candidate has so many votes that he will win whatever happens with the remaining votes. Being able to determine if a candidate must win is useful as we can stop eliciting preferences.

In addition to uncertainty about the agents' preferences, we may have uncertainty about how the voting rule will be applied. For instance, in sequential majority voting (sometimes called the "Cup" or "tournament" rule), preferences are aggregated by a sequence of pairwise comparisons (also called an "agenda"). The order of these comparisons may not yet be fixed or may not be known. Nevertheless, we may still be able to determine information about the outcome. For example, it may be that one candidate cannot win however the voting rule is applied. This is useful, for example,

if we want to know if the chair can manipulate the election to make their favoured candidate win.

In this paper we study the computational complexity of determining the possible and necessary winners in sequential majority voting with weighted agents, when preferences may be incomplete and we may not know the agenda. We show that determining if a candidate must win in every agenda is polynomial. However, determining if a candidate can win in at least one completion of the unrevealed preferences and at least one agenda is NP-hard. This problem remains hard if the agenda is balanced.

Background

Preferences and profiles. We assume that each agent's preferences are specified by a (possibly incomplete) total order (TO) (that is, by an asymmetric, irreflexive and transitive order) over a set of candidates (denoted by Ω). Given two candidates, say $A, B \in \Omega$, an agent specifies exactly one of the following: $A < B$, $A > B$, or $A?B$, where $A?B$ means that the relation between A and B has not yet been revealed. We assume that an agent's preferences are transitively closed. That is, if they declare $A > B$, and $B > C$ then they also have $A > C$. A *weighted profile* is a sequence of total orders describing the preferences for n agents, each of which has a given weight. A weighted profile is *incomplete* if one or more of the preference relations is incomplete. For simplicity, we assume that the sum of the weights of agents is odd. An (incomplete) *unweighted profile* is one in which each agent has weight 1. From any weighted profile P, we can build the *corresponding unweighted profile* P' by replacing every ordering expressed by an agent with weight k_i by k_i agents with weight 1 expressing the identical ordering.

Majority graphs. Given an (incomplete) weighted profile P, the *majority graph $M(P)$* induced by P is the directed graph whose set of vertices is Ω and where an edge from A to B (denoted by $A >_m B$) denotes a strict weighted majority of voters who prefer A to B. A majority graph is said to be complete if, for any two vertices, there is a directed edge between them. Notice that, if P is incomplete, $M(P)$ may be incomplete as well. Moreover, if $M(P)$ is incomplete, the set of all complete majority graphs extending $M(P)$ corresponds to a (possibly proper) superset of the set

of complete majority graphs induced by all possible completions of P. This is due to correlations between votes which might prevent a given graph from being implementable.

Example 1 *Consider the incomplete weighted profile P composed by three agents a_1, a_2 and a_3 with weights resp. 1, 2 and 2 where a_1 states $A > B > C$, a_2 states $B > A$, $A?C$, $B?C$ and a_3 states $A > B$, $A?C$, $C > B$. The majority graph induced by P, called $M(P)$, is the graph with three nodes A, B and C and only one edge $A >_m B$.*

Sequential majority voting. Given a set of candidates, the sequential majority voting rule is defined by a binary tree (also called an *agenda*) with one candidate per leaf. Each internal node represents the candidate that wins the pairwise election between the node's children. The winner of every pairwise election is computed by the weighted majority rule, where A beats B iff there is a weighted majority of votes stating $A > B$. The candidate at the root of the agenda is declared the overall winner. Given a complete profile, candidates which win whatever the agenda are called *Condorcet winners*.

Winners from majority graphs. Four types of potential winner have been defined (Lang *et al.* 2007). Given and an incomplete majority graph G induced by an incomplete profile P, consider a candidate A. Then

- A is a *weak Condorcet winner* for G ($A \in WC(G)$) iff there is a completion of G such that A wins in every agenda;

- A is a *strong Condorcet winner* for G ($A \in SC(G)$) iff for every completion of G, A wins in every agenda;

- A is a *weak possible winner* for G ($A \in WP(G)$) iff there exists a completion of G and an agenda for which A wins;

- A is a *strong possible winner* for G ($A \in SP(G)$)) iff for every completion of G there is an agenda for which A wins.

When the majority graph is complete, strong and weak Condorcet winners coincide (that is, $SC(G) = WC(G)$). Similarly, strong and weak possible winners coincide in this case (that is, $SP(G) = WP(G)$). In (Lang *et al.* 2007), it is proved that $WP(G)$, $SP(G)$, $WC(G)$, and $SC(G)$ can all be computed in polynomial time.

Profiles vs. majority graphs

These notions of possible and Condorcet winner are based on an incomplete majority graph. It is, however, often more useful and meaningful to start directly from the incomplete profile inducing the majority graph. Given an incomplete profile, there can be more completions of its induced majority graphs than majority graphs induced by completing the profile. The problem is that the incomplete majority graph throws away information about how individual agents have voted. For example, we lose information about correlations between votes. Such correlations may prevent a candidate from being able to win.

Example 2 *Consider an incomplete profile P with just one agent and three candidates (A, B, and C), where the agent*

declares only $A > B$. Then the induced majority graph $M(P)$ has only one arc from A to B. In this situation, B is a weak possible winner (that is, $B \in WP(M(P))$), since there is a completion of the majority graph (where B beats C and C beats A), and an agenda where B wins (we first compare A with C, C wins, and then C with B, where B wins). However, there is no way to complete profile P and set up the agenda so B wins. It is therefore rather misleading to consider B as a potential winner.

Hence, unlike (Lang *et al.* 2007), we will define possible and Condorcet winners starting directly from profiles, rather than the induced majority graphs.

Weighted votes

As in (Conitzer & Sandholm 2002a), we will consider weighted votes. Although human elections are often unweighted, the addition of weights makes voting schemes more general. Weighted voting systems are used in a number of real-world settings like shareholder meetings, and elected assemblies. Weights are useful in multiagent systems where we have different types of agents. Weights are also interesting from a computational perspective.

First, as we argue here, computing the weak/strong possible/Condorcet winners with unweighted votes is always polynomial. If there a bounded number of candidates, there are only a polynomial number of different ways to complete the profile or majority graph. There are also only a polynomial number of different agendas. All the possibilities can therefore be tested in polynomial time. On the other hand, adding weights to the votes may introduce computational complexity. For example, as we will show later, computing weak possible winners becomes NP-hard when we add weights. Second, as argued in (Conitzer & Sandholm 2002a) for manipulation, if it is hard to compute possible winners with weighted votes, it will also be hard to compute the probability of winning in the unweighted case when there is uncertainty about how the votes have been cast. Thus, the weighted case informs us about the unweighted case in the presence of uncertainty about the votes.

Possible and Condorcet winners from profiles

Given an incomplete weighted profile, we introduce the following notions of weak/strong possible/Condorcet winners.

Definition 1 *Let P be an incomplete weighted profile and A a candidate.*

- *A is a* weak Condorcet winner *for P ($A \in WC(P)$) iff there is a completion of P such that A is a winner for all agendas;*

- *A is a* strong Condorcet winner *for P ($A \in SC(P)$) iff for every completion of P, and for every agenda, A is a winner;*

- *A is a* weak possible winner *for P ($A \in WP(P)$) iff there exists a completion of P and an agenda for which A wins;*

- *A is a* strong possible winner *for P ($A \in SP(P)$) iff for every completion of P there is an agenda for which A wins;*

It is easy to see that, when the profile is complete, strong and weak Condorcet winners coincide. The same holds also for strong and weak possible winners.

We now make a comparison between the notions of winners defined in (Lang *et al.* 2007) and those defined above in this paper. Since in (Lang *et al.* 2007) weights were not considered, we first consider unweighted profiles. Given an incomplete unweighted profile P and the incomplete majority graph G induced by P, that is, $G = M(P)$, we know that the completions of G are a (possibly proper) superset of the set of complete majority graphs induced by all possible completions of P. This simple observation leads to the following results.

Theorem 1 *Given an incomplete unweighted profile P,*

- $WP(M(P)) \supseteq WP(P)$;
- $SP(M(P)) \subseteq SP(P)$;
- $WC(M(P)) = WC(P)$;
- $SC(M(P)) = SC(P)$.

Proof:

- $WP(M(P)) \supseteq WP(P)$.
  If a candidate A belongs to $WP(P)$, there is a completion of P, say P', and an agenda, such that A wins. Thus $A \in WP(G')$ where G' is the complete majority graph induced by P'. Since G' is one of all the possible completions of $M(P)$), then $A \in WP(M(P))$.

- $SP(M(P)) \subseteq SP(P)$.
  If a candidate is a possible winner for every completion of G, it is also a possible winner for the majority graphs induced by the completions of P, since they are a subset of the set of all the completions of $M(P)$.

- $WC(M(P)) = WC(P)$.
  The same reasoning used in the first part of this proof can be used here to show that $WC(M(P)) \supseteq WC(P)$. We can also prove that $WC(M(P)) \subseteq WC(P)$. In fact, if a candidate A belongs to $WC(M(P))$, then there must be one or more completions of the majority graph where A has only outgoing edges. Among such completions, there is for sure one which derives from a completion of the profile in which all $A?C$ become $A > C$ (for all C). Thus, setting this is sufficient to make A a weak Condorcet winner and does not give any transitivity problems in the profile.

- $SC(M(P)) = SC(P)$.
  The same reasoning used in the second part of this proof can be used here to show that $SC(M(P)) \subseteq SC(P)$. We can also prove that $SC(M(P)) \supseteq SC(P)$. In fact, if a candidate belongs to $SC(P)$, then it is a Condorcet winner, i.e., it beats every other candidate, for every completion of P. Thus it must beat every other candidate in the certain part. Thus in the (possibly incomplete) majority graph $M(P)$ induced by P there are only outgoing edges from this candidate, and so this candidate must belong to $SC(M(P))$. $\square$

Notice that there are cases in which the subset relation $WP(M(P)) \supseteq WP(P)$ is strict. In fact, a candidate can be a possible winner for a completion of $M(P)$ which is not induced by any completion of P, as shown previously in Example 2.

We next consider weighted profiles. Although weighted profiles were not considered in (Lang *et al.* 2007), the same notions defined in that paper can be given also for majority graphs induced by weighted profiles. We will now show that the same results as in Theorem 1 hold also in this more general setting. To do this, we first show that, given an incomplete weighted profile P and the corresponding unweighted profile P', $SC(P) = SC(P')$ and $WC(P) = WC(P')$.

Theorem 2 *Given an incomplete weighted profile P, let P' the corresponding unweighted profile obtained from P. Then*

- $M(P) = M(P')$;
- $SC(P) = SC(P')$;
- $WC(P) = WC(P')$.

Proof:

- $M(P) = M(P')$.
  We show that $\forall A, B \in \Omega$, $A >_m B$ in $M(P)$ iff $A >_m B$ in $M(P')$ and $A?_m B$ (i.e., $A \not>_m B$ and $B \not>_m A$) in $M(P)$ iff $A?_m B$ in $M(P')$. Given candidates A and B, $A >_m B$ in $M(P)$ iff there are j agents with total weights $T_j > (\sum_{i=1}^{n} k_i)/2$ stating $A > B$ in P. This happens iff there are T_j agents with weights 1 stating $A > B$ in P', i.e., there is a majority of agents stating $A > B$ in P'. This means that $A >_m B$ in $M(P')$.

  Given candidates A and B, $A?_m B$ in $M(P)$ iff there is no weighted majority of agents in P all stating $A > B$ or all stating $B > A$. This means that in P' there is no majority of agents all stating $A > B$ or all stating $B > A$. Thus $A \not>_m B$ and $B \not>_m A$ in $M(P')$, i.e., $A?_m B$ in $M(P')$.

- $SC(P) = SC(P')$.
  ($\Leftarrow$) It follows from the fact that the set of completions of P' is a superset of the set of the completions of P.
  ($\Rightarrow$) Assume that $A \notin SC(P')$. Then A has not $m - 1$ outgoing edges (where $m = |\Omega|$) in $M(P')$ (Lang *et al.* 2007). Hence, since $M(P) = M(P')$, A has not $m - 1$ outgoing edges in $M(P)$. Hence there is a candidate B s.t. $B >_m A$ or $B?_m A$ in $M(P)$. If $B >_m A$ in $M(P)$, then for every completion of P we have $B > A$, and thus A cannot win in every agenda. If $B?_m A$ in $M(P)$, then there exists a completion of P where we replace every $A?B$ with $B > A$, where A may not win. Thus it is not true that A wins for every completion and for every agenda.

- $WC(P) = WC(P')$.
  ($\Rightarrow$) It follows from the fact that the set of completions of P' is a superset of the set of completions of P.
  ($\Leftarrow$) Assume that $A \in WC(P')$. Then A has no ingoing edges in $M(P')$ (Lang *et al.* 2007). Hence, since $M(P') = M(P)$, A has no ingoing edges in $M(P')$. Thus if we replace, for every B, $A?B$ in P with $A > B$, there we obtain a completion of P where A wins for every agenda. Thus $A \in WC(P)$. $\square$

We are now ready to compare the notions of winners in the weighted case.

Theorem 3 *Given an incomplete weighted profile P, we have:*

- $WP(M(P)) \supseteq WP(P)$;
- $SP(M(P)) \subseteq SP(P)$;
- $SC(M(P)) = SC(P)$;
- $WC(M(P)) = WC(P)$.

Proof: Let P' be the corresponding unweighted profile obtained from P. Since the set of the completions of P is a subset of the set of completions of P', we have that $WP(P) \subseteq WP(P')$ and $SP(P) \supseteq SP(P')$. Now, since $M(P) = M(P')$ by Theorem 2, and since $SP(G)$ and $WP(G)$ depend only on the majority graph G under consideration, we have that $WP(M(P)) \supseteq WP(P)$ and $SP(M(P)) \subseteq SP(P)$. To prove that $SC(P) = SC(M(P))$, we may notice that $SC(P) = SC(P')$ by Theorem 2, $SC(P') = SC(M(P'))$ by Theorem 1, and $SC(M(P') = SC(M(P))$ by Theorem 2 and by the fact that $SC(G)$ depends only on the majority graph G considered. The same reasoning allows us to conclude that $WC(P) = WC(M(P))$. □

Complexity of determining winners

We now turn our attention to study the complexity of determining possible and Condorcet winners. We start by showing that computing weak and strong Condorcet winners is polynomial in the number of agents and candidates.

Theorem 4 *Given an incomplete weighted profile P, the sets $WC(P)$ and $SC(P)$ are polynomial to compute.*

Proof: By Theorem 3, $WC(P) = WC(M(P))$ and $SC(P) = SC(M(P))$. Moreover, by Theorem 2 we know that $M(P) = M(P')$, where P' is the corresponding unweighted profile obtained from P. Thus $WC(P) = WC(M(P'))$ and $SC(P) = SC(M(P'))$. In (Lang *et al.* 2007) the authors show that, given any majority graph G obtained from an unweighted profile, it is polynomial to compute $WC(G)$ and $SC(G)$. Thus it is polynomial to compute $WC(M(P'))$ and $SC(M(P'))$. □

We now consider weak possible winners. We show that, computing weak possible winners is intractable in general.

Theorem 5 *Given an incomplete weighted profile P with 3 or more candidates, deciding if a candidate is a weak possible winner for P is NP-complete.*

Proof: We give a reduction from the number partitioning problem. We have a bag of integers, k_i with sum $2k$ and we wish to decide if they can be partitioned into two bags, each with sum k. We want to show that a candidate B is a weak possible winner if and only if such a partition exists. We construct an incomplete profile over three candidates (A, B, and C) as follows. We have 1 vote for $B > C > A$ of weight 1, 1 vote $B > A > C$ of weight $2k - 1$, and 1 vote $C > B > A$ of weight $2k - 1$. At this point, the weight of votes such that B is ahead of A is $4k - 1$, the weight of votes such that B is ahead of C is 1, and the weight of votes

such that C is ahead of A is 1. We also have, for each k_i, a partially specified vote of weight $2k_i$ in which we know just that $A > B$. As the total weight of these partially specified votes is $4k$, we are sure A beats B in the final result by 1 vote. The partially specified votes can be completed to make $A > B$, $B > C$, and $C > A$ iff there is a partition of size k. Suppose there is such a partition. Then let the votes in one bag be $A > B > C$ and the votes in the other be $C > A > B$. Then, A beats B, B beats C and C beats A, all by 1 vote. On the other hand, suppose there is a way to cast the votes to give the result A beats B, B beats C and C beats A. All the uncast votes rank A above B. In addition, at least half the weight of votes must rank B above C, and at least half the weight of votes must rank C above A. Since A is above B, C cannot be both above A and below B. Thus precisely half the weight of votes ranks C above A and half ranks B above C. Thus we have a partition of equal weight. Now, as A beats B, B is a possible winner for the considered completion of the profile iff B beats C and C beats A. Thus, B is a weak possible winner iff there is a partition of size k. □

Note that computing weak possible winners from an incomplete majority graph is polynomial (Lang *et al.* 2007). Thus, adding weights to the votes and computing weak possible winners from the incomplete profile instead of the majority graph makes the problem intractable. On the other, adding weights to the votes did not make weak and strong Condorcet winners hard to compute.

Weak fair possible winners

In (Lang *et al.* 2007) the notion of fair winner in sequential majority voting is introduced. In particular: given a complete profile P, a candidate A is said to be a *fair possible winner* for P iff there is a balanced agenda in which A wins. If the number of candidates is a power of two, a balanced agenda is a balanced binary tree in which every candidate needs to win the same number of pairwise comparisons to win overall. If the number of candidates is not a power of two, each candidate has at most one bye. The motivation to consider fair possible winners is to avoid situations in which weak candidates, that can beat only a small number of candidates, end up winners of the election. In (Lang *et al.* 2007) it has been shown that testing whether a candidate is a fair possible winner over weighted majority graphs is NP-hard.

We consider again this notion of fairness, applied to the new notions of possible winner. In particular, we say that: given an incomplete weighted profile P, A is a *fair weak possible* (FWP) winner for P iff there exists a completion of P and a balanced agenda such that A wins. Given the proof of Theorem 5, it is easy to see that determining such winners is difficult. However, this result cannot be derived from the analogous one in (Lang *et al.* 2007), since we don't have weights in the majority graph.

Theorem 6 *Given an incomplete weighted profile P with 3 or more candidates, deciding if a candidate is a fair weak possible winners for P is NP-complete.*

Proof: It follows from the proof of Theorem 5, since every agenda with 3 candidates is represented by a balanced

agenda. □

Related work

Contizer and Sandholm prove that deciding if preference elicitation is over (that is, determining if the remaining votes can be cast so a given candidate does not win) is NP-hard for the STV rule (Conitzer & Sandholm 2002b). For other common voting rules like plurality, Borda and the sequential majority rule, they show that it is polynomial to decide if preference elicitation is over.

Konczak and Lang show that it is polynomial to compute possible and necessary winners for positional scoring voting rules like the Borda and plurality rule, as well as for a non-positional rule like Condorcet (Konczak & Lang 2005). They argue that elicitation is over when the set of possible winners contains just the necessary winner. They also argue that if computing possible (resp., Condorcet) winners is polynomial, then constructive (resp., destructive) manipulation of the election is polynomial.

Pini *et al.* prove that, for the STV rule, computing the possible and necessary winners is NP-hard (Pini *et al.* 2007). In fact, they show it is NP-hard even to approximate these sets within some constant factor in size. They also give a preference elicitation procedure which focuses just on the set of possible winners. Lang *et al.* consider determining the winner for the sequential majority voting rule in the presence of uncertainty about the votes and agenda (Lang *et al.* 2007). As mentioned earlier, the major difference is that this work starts from incomplete majority graphs whilst we start from incomplete profiles. The incomplete majority graph throws away information about the individual votes. For this reason, it may suggest candidates can win when they cannot.

Finally, Conitzer and Sandholm show that, if the agenda is fixed, determining the weak winners is polynomial, but randomizing the agenda makes deciding the probability that a candidate wins (and thus manipulation) NP-hard (Conitzer & Sandholm 2002a). They also prove that constructive manipulation is intractable for the Borda, Copeland, Maximin and STV rules using weighted votes even with a small number of candidates. However, all of these rules are polynomial to manipulate destructively except STV.

Conclusions

We have considered agents combining preferences using the sequential majority rule. We have studied the situation where agents may not have revealed all their preferences (either because we are still eliciting preferences or because of issues like privacy or communication cost). We have also considered uncertainty in how the voting rule is applied. In these settings, we have studied the computational complexity of computing whether a candidate must or can win. The following table summarizes the complexity results discussed in this paper. The table has one cell for each of the several notions of winners. Each row considers the same notion based either on an incomplete majority graph, on an incomplete profile. The new results are those appearing in the second column of the table. The results in the first column were presented in (Lang *et al.* 2007).

	maj.graph	profile
WP	P	NP-hard
WC	P	P
SC	P	P
FWP	NP-hard (with weighted maj.graph)	NP-hard

Notice that the complexity of determining fair weak possible winners, with or without weights for the agents, was still open before the NP-hardness result given here. In fact, the only existing result was for unweighted agents and majority graphs in which the edges are labelled with weights. Such labels could represent, for example, the amount of disagreement between the agents. In this paper, edges in majority graphs are not labelled with weights, but are simply directed according to the majority weight of votes.

These results are useful in determining if preference elicitation is over. They are also useful to determine how difficult it is for the chair to control the election. As future work we want to determine the computational complexity of finding strong possible winners from incomplete weighted profiles. For incomplete majority graphs such a computation is polynomial (Lang *et al.* 2007), we want to check if there is the same complexity also considering incomplete weighted profiles. Moreover, we intend to investigate the complexity of determining winners from incomplete weighted profiles when the agenda is fixed. In such a case we can define new notions of winners: weak winners, i.e., those candidates that win in the fixed agenda for at least a completion of the incomplete profile and strong winners, i.e., those candidates that win in the fixed agenda for every completion of the incomplete profile. Another interesting direction for future work is deciding which candidate or candidates are most likely to win. This is related to probabilistic approaches to voting theory. Another interesting direction is to study other forms of uncertainty in the application of the voting rule (e.g. if we have uncertain weights in a scoring rule, or if the chair can choose between a certain set of voting rules).

References

Conitzer, V., and Sandholm, T. 2002a. Complexity of manipulating an election with few candidates. In *Proceedings of AAAI-02*, 314–319.

Conitzer, V., and Sandholm, T. 2002b. Vote elicitation: complexity and strategy-proofness. In *Proceedings of AAAI-02*, 392–397.

Konczak, K., and Lang, J. 2005. Voting procedures with incomplete preferences. In *Proc. IJCAI-05 Multidisciplinary Workshop on Advances in Preference Handling*.

Lang, J.; Pini, M. S.; Rossi, F.; Venable, K. B.; and Walsh, T. 2007. Winner determination in sequential majority voting. In *Proceedings of IJCAI-07*.

Pini, M. S.; Rossi, F.; Venable, K. B.; and Walsh, T. 2007. Incompleteness and incomparability in preference aggregation. In *Proceedings of IJCAI-07*.