Voting procedures with incomplete preferences

Kathrin Konczak
Institut für Informatik, Universität Potsdam
Postfach 90 03 27, D-14439 Potsdam
konczak@cs.uni-potsdam.de

Jerome Lang
IRIT- Univ. Paul Sabatier
Toulouse, France
lang@irit.fr

Abstract

We extend the application of a voting procedure (usually defined on complete preference relations over candidates) when the voters' preferences consist of partial orders. We define possible (resp. necessary) winners for a given partial preference profile $\mathcal{R}$ with respect to a given voting procedure as the candidates being the winners in some (resp. all) of the complete extensions of $\mathcal{R}$. We show that, although the computation of possible and necessary winners may be hard in general case, it is polynomial for the family of positional scoring procedures. We show that the possible and necessary Condorcet winners for a partial preference profile can be computed in polynomial time as well. Lastly, we point out connections to vote manipulation and elicitation.

1 Introduction

Automated group decision making is an important issue in AI: autonomous agents often have to agree on a common decision, and may for this reason apply *voting procedures*, which is one of the most common ways of making a collective choice. Voting procedures, studied extensively by social choice theorists from the normative point of view, have been recently studied from the computational point of view:

- while winner determination is easy with most usual voting procedures (at least when the number of candidates is small), a few of them are hard, and their complexity and their practical computation have been investigated in [Rothe *et al.*, 2003; Hemaspaandra *et al.*, 2004; Davenport and Kalagnanam, 2004].
- some works focus on sets of candidates with a combinatorial structure and investigate compact representation issues [Lang, 2004; Rossi *et al.*, 2004] and the complexity of voting procedures applied in such domains [Lang, 2004].
- even when computing the outcome a voting procedure is easy, it might be the case that determining whether there is a successful manipulation for a coalition of voters is hard; the complexity of this problem is studied in [Conitzer and Sandholm, 2002a; Conitzer *et al.*, 2003; Conitzer and Sandholm, 2003].
- elicitation issues and partial winner determination when the preference profile is not fully known has been studied in [Conitzer and Sandholm, 2002b].

In this paper we focus on the last of these issues, which raises the question of the application of a voting procedure when the voters' preferences are incomplete. Let $X = \{x_1, \ldots, x_m\}$ be a finite set of candidates and $I = \{1, \ldots, n\}$ be a finite set of *voters* ; a collective preference profile is a collection of partial preference relations $\mathcal{R} = \langle R_1, \ldots, R_n \rangle$ on X (formal details follow in Section 3). Winner determination under incomplete preference consists in applying (in some sense that we make precise later on) a voting procedure over such a collective preference profile. This is particularly relevant in the following situations:

- some voters have expressed their preference profile and some others have not yet done it; in that case, the collective preference profile is a collection consisting of n_1 complete preference relations and $n - n_1$ empty preference relations.
- all voters have expressed their preferences on a given subset of candidates, and now new candidates are introduced, about which the voters' preferences are unknown.
- voters are allowed to express their preferences in an incremental way: they left some comparisons between candidates unspecified, because either they *don't know* or they *don't want* to compare some candidates (we comment further on the various possible interpretations of incomplete preferences).
- preferences have been only partially elicited and/or are expressed in a language for compact preference representation such as CP-nets [Boutilier *et al.*, 2004] which induce partial preference relations in the general case.

In all these cases it may be worth having an idea of the possible outcomes of the vote without waiting for the preferences to be complete (which sometimes will never happen, as in "refuse to compare" case mentioned above). In some cases we may conclude that the preferences known so far, although incomplete, are informative enough so that the outcome of the vote can be determined; if this is not the case, we may compute a set of candidates that may win the vote after the preferences have become complete, thus giving the voters an opportunity to focus on these candidates and forget about the others. Lastly, similarly as in [Conitzer and Sandholm, 2002b], we may determine from these incomplete preferences which preferences should be elicited from whom so as to be able to compute the winner.

In Section 2 we give some basic background on voting procedures. In Section 3, after briefly discussing three different ways of applying voting procedures (tailored for complete preferences) to incomplete preferences, we will focus on one of these ways, that we think is more suited to the case where incompleteness corresponds to an incomplete knowledge of the voters' preferences; we will then introduce the natural notions of necessary and possible winners for a given partial preference profile and a voting procedure. In Section 4 we show that in the case of positional scoring procedures, possible and necessary winners can be computed in polynomial time by a very simple algorithm. In Section 5 we investigate the notion of possible and necessary Condorcet winners and show that they can be computed as well in polynomial time. Section 6 considers related issues such as vote manipulation and elicitation.

2 Background

2.1 Preference relations

We start by giving some terminology and notations about preference relations and voting procedures.

An *order* R on X is a reflexive, transitive and antisymmetric relation on X (recall that R is antisymmetric if and only if for all $x, y \in X$, $R(x, y)$ and $R(y, x)$ implies $x = y$.) $R(x, y)$ is also denoted by $x \geq y$. $\succ_R$ denotes the strict relation induced from R, defined by $x \succ_R x'$ if and only if $x \succeq_R x'$ and not $(x' \succeq_R x)$[1]. An order R is *linear* (or *complete*) if and only if $R(x, x')$ or $R(x', x)$ holds for all $x, x' \in X$. Let R, R' be two orders on X. R' *extends* R if and only if $R \subseteq R'$, that is, $R(x, x')$ implies $R'(x, x')$ for all $x, x' \in X$. Let R be an order. A linear order T is an *complete extension* of R if and only if T extends R. $Ext(R)$ denotes the set of all complete extensions of R.

2.2 Voting procedures

Let $X = \{x_1, \ldots, x_m\}$ be a finite set of *candidates* and $I = \{1, \ldots, n\}$ a finite set of *voters*. A *complete preference profile* is a collection $\mathcal{T} = \langle T_1, \ldots, T_n \rangle$ of linear orders on X (where T_i represents the preferences of voter i). A *voting procedure* F maps every complete preference profile $\mathcal{T}$ to a nonempty subset of X: $F(\mathcal{T})$ denotes the set of winners of $\mathcal{T}$ w.r.t. F. Note, importantly, that the outcome $F(\mathcal{T})$ of a voting procedure is always nonempty, i.e., the outcome of the procedure is defined for all preference profiles.

Examples of voting procedures are considered in Sections 4 and 5. Among the many voting procedures that exist in the literature (for an extensive presentation see for instance [Brams and Fishburn, 2003]), some require preference profiles to be linear orders and some allow more generally preferences to be weak orders (where antisymmetry is not required, which implies that a voter can express an indifference between two candidates). However, for the sake of simplicity, in this paper we assume that all preference relations considered are antisymmetric. This is not a real loss of generality, as most of our definitions and results would extend to the case where indifference is allowed (see Section 7).

Even if some voting procedures work on linear orders and some on weak orders, a common point of all procedures is that they apply to *complete* preference relations: in other words, they are not tailored for dealing with *incomparability*. In Section 1 we argued towards taking possible incomparabilities into account. The question now is how F should be extended when we have only a partial knowledge of the preferences of the voters – in other terms, how should F be defined when the input is a collection of *orders* rather than a collection of *linear orders*. This issue is investigated in the next section.

3 Voting procedures with incomplete preferences: definitions

3.1 Extending voting procedures to incomplete preferences

A *voting problem under incomplete preferences* is composed of a finite set of *candidates* $X = \{x_1, \ldots, x_m\}$, a finite set of *voters* $I = \{1, \ldots, n\}$, and for each i, an order R_i on X denoting the *individual preference profile* of voter i. The collection of orders $\mathcal{R} = \langle R_1, \ldots, R_n \rangle$ will be called a *(collective) preference profile*. $\mathcal{R}$ is said to be complete if and only if R_i is complete for each $0 \leq i \leq n$. In the rest of the paper, R_i is often denoted as $\succeq_{R_i}$ or as $\succeq_i$: thus, we write indifferently $R_i(x, y)$, $x \succeq_{R_i} y$, or $x \succeq_i y$.

The notion of complete extension is generalized from individual to collective preference profiles in a natural way:

$$Ext(\mathcal{R}) = Ext(R_1) \times \ldots \times Ext(R_n)$$

There are at least two interpretations for incomplete preferences: *intrinsic* incompleteness, where the voter refuses to compare some alternatives, or *epistemic* incompleteness, where the voter has a complete preference but it is only partially known at the time the voting procedure has to be applied. These different interpretations lead to different ways of extending voting procedures to partial preferences, as discussed further. Here are three possible ways that can be followed, where F is a given voting procedure defined for complete preference relations: (1) apply F to all complete extensions of the preference relations and gather the results; (2) select a subset of those complete extensions (ideally a singleton) using some completion process, apply F to these and gather the results; [2]. (3) rewrite directly the definition of F so that it applies more generally to partial preference relations (obviously, this extension of F must coincide with F on complete preference profiles).

[1] Or equivalently, since R is antisymmetric: $x \succ_R x'$ if and only if $x \succeq_R x'$ and $x' \neq x$).

[2] This completion process may consist in letting candidates gravitate towards preference such as in [Boutilier, 1994] or towards indifference such as in [Tan and Pearl, 1994].

In the rest of the paper we explore only the first of these three ways, which looks the most natural of all three ways; furthermore it seems to be more suited to epistemic incompleteness of preference (see Section 6).

3.2 Possible and necessary winners

For applying a voting procedure to all complete extensions of a partial preference profile we define upper and lower bounds for winners.

Definition 1 *Let F be a voting procedure on X and $\mathcal{R}$ a (possibly incomplete) preference profile.*

- *$x \in X$ is a* necessary winner *for $\mathcal{R}$ (w.r.t. F) if and only if for all $\mathcal{T} \in Ext(\mathcal{R})$ we have $x \in F(\mathcal{T})$.*
- *$x \in X$ is a* possible winner *for $\mathcal{R}$ (w.r.t. F) if and only if there exists a $\mathcal{T} \in Ext(\mathcal{R})$ such that $x \in F(\mathcal{T})$.*

A necessary winner for $\mathcal{R}$ is thus a candidate which wins in all complete extensions of $\mathcal{R}$ and a possible winner wins in at least one complete extension of $\mathcal{R}$. Hence, necessary winners constitute an upper bound and possible winners a lower bound for winners of a partial preference profile. We denote by $NW_F(X)$ (respectively $PW_F(X)$) the set of necessary (respectively possible) winners for $\mathcal{R}$ w.r.t. F. Clearly, the following properties hold, for any voting procedure F:

- for all $\mathcal{R}$, $NW_F(\mathcal{R}) \subseteq PW_F(\mathcal{R})$;
- for all $\mathcal{R}, \mathcal{R}'$ such that $\mathcal{R} \subseteq \mathcal{R}'$, $PW_F(\mathcal{R}') \subseteq PW_F(\mathcal{R})$ and $NW_F(\mathcal{R}') \subseteq NW_F(\mathcal{R})$.

Note also that $NW_F(X)$ can be empty, but not $PW_F(X)$. Whenever $\mathcal{R}$ is a complete preference profile, possible and necessary winners coincide.

The rest of these notes try to evaluate the difficulty of applying some well-known voting procedures to partial preference relations, by assessing the computational complexity of the problems and by giving explicit algorithms for computing possible and necessary winners.

Since there are, in the general case, exponentially many extensions of a partial preference profile, nothing guarantees that computing possible and necessary winners can be done in polynomial time, even if the voting procedure F is polynomially computable. All we can say is that, provided that F is polynomially computable:

- Determining whether $x \in PW_F(\mathcal{R})$ is in NP.
- Determining whether $x \in NW_F(\mathcal{R})$ is in coNP.

The question is now: are there any voting procedures such that necessary and possible winners can still be determined in polynomial time? We answer this question positively in the next two Sections.

4 Positional scoring procedures

A *positional scoring procedure* is defined from a *scoring vector* , that is, a vector $\vec{s} = (s_1, \ldots, s_m)$ of integers such that $s_1 \geq s_2 \geq \ldots \geq s_m$ and $s_1 > s_m$. Let $\mathcal{T} = \langle T_1, \ldots, T_n \rangle$ be a complete preference profile. For every $x \in X$ and every $i \in I$, let $r(T_i, x) = \#\{y \mid y >_{T_i} x\} + 1$ be the rank of x in the complete order T_i; then

$$S(x, \mathcal{T}) = \sum_{i=1}^{n} s_{r(T_i,x)}$$

Lastly, the positional scoring rule $F_{\vec{s}}$ associated with a scoring vector $\vec{s}$ is defined by its set $F_{\vec{s}}(\mathcal{T}) = \{x \mid S(x, \mathcal{T}) \text{ is maximal }\}$, that is, the set of winning candidates for $\mathcal{T}$ with respect to $F_{\vec{s}}$ is the set of candidates in X maximizing $S(., \mathcal{T})$.

Here are well-known examples of positional scoring procedures:

- the *Borda* procedure is the defined from the scoring vector $s_k = m - k$ for all $k = 1, \ldots, m$;
- the *plurality* procedure is defined from the scoring vector $s_1 = 1$, and $s_k = 0$ for all $k > 1$.

The question now is, how hard is it to determine whether x is a necessary or a possible winner for $\mathcal{R}$ w.r.t. a scoring procedure F?

For this, let us define the *minimal* (resp. *maximal*) *rank* of a candidate x for a (partial) order R as the lowest[3] (resp. highest) possible rank of x obtained when considering all complete extensions of R, that is,

$$rank_R^{min}(x) = \min_{T \in Ext(R)} r(T, x)$$
$$rank_R^{max}(x) = \max_{T \in Ext(R)} r(T, x)$$

where $r(T, x)$ is the rank of x in the complete order T, which is a complete extension of R.

These bounds are actually much easier to compute than what their definition suggests. For some voter, the minimal rank of x is determined by the number of candidates which are higher ranked in the order and the maximal rank of x is determined by the number of lower ranked candidates.

Proposition 1 *Let R be a (partial) strict order. Then,*

$$rank_R^{min}(x) = \#\{y \mid y >_R x\} + 1 \textit{ and}$$
$$rank_R^{max}(x) = m - \#\{y \mid x >_R y\}$$

Proof: $rank_R^{min}(x)$**:** Let $R_x^{\uparrow}$ be the following extension of R:

$$(\uparrow)\ R_x^{\uparrow} = R \cup \{(x, z) \mid z \neq x, not\ (z >_R x)\}$$

That is, all candidates which are not initially strictly preferred to x are now less preferred: for all $z \neq x$, $x > z$ holds in $R_x^{\uparrow}$ as soon as $(z >_R x)\}$ does not hold, that is, as soon as it is possible to enforce $x > z$. Note also that it can be easily checked that $R_x^{\uparrow}$ is transitive. Next, we show that $R_x^{\uparrow}$ has at least one complete extension. For this, let $\langle z_i \rangle_{i \in L}$ be an enumeration of $\{z \mid z \neq x, not(z >_R x)\}$ such that $i < j$ if $z_i >_R z_j$ holds for all $i, j \in L$. Analogously, let $\langle y_i \rangle_{i \in K}$ be an enumeration of $\{y \mid y >_R x\}$ such that $i < j$ if $y_i >_R y_j$ holds for all $i, j \in K$. Then, let $\langle r_i \rangle_{i \in J}$ be an enumeration of $\mathcal{X}$ such that $\langle r_i \rangle_{i \in J} = \langle \langle y_i \rangle_{i \in K}, x, \langle z_i \rangle_{i \in L} \rangle$ Then, R', defined as $r_i \geq_{R'} r_j$ iff $i \leq j$, is an complete extension of $R_x^{\uparrow}$.

For all complete extensions of $R_x^{\uparrow}$ we have that x has the rank $\#\{y \mid y >_R x\} + 1$ since all other candidates are ranked lower than x. For all complete extensions T which do not satisfy $(\uparrow)$ we have that x has at least the rank $\#\{y \mid y >_R x\} + 2$ since there is at least one other candidate who is additionally higher ranked than x. Hence, we have $rank_R^{min}(x) = \#\{y \mid y >_R x\} + 1$.

[3] Recall that the lower its rank, the more preferred a candidate.

$rank_{\mathcal{R}}^{max}(x)$**:** Proof is similar by taking ($\downarrow$) $R_x^{\downarrow} = R \cup \{(z,x) \mid z \neq x, not\ (x >_R z)\}$ ■

Next, necessary and possible winners can be computed by consider the best and the worst case for values of scoring functions.

Proposition 2 *Let $\mathcal{R}$ be a preference profile, where each R_i is a (partial) order, F_s be a positional voting procedure, and*

$$S_{\mathcal{R}}^{min}(x) = \sum_{i=1}^{n} s_{rank_{R_i}^{max}(x)}$$
$$S_{\mathcal{R}}^{max}(x) = \sum_{i=1}^{n} s_{rank_{R_i}^{min}(x)}$$

Then,
(1) x is a necessary winner for $\mathcal{R}$ w.r.t. F_s if and only if $S_{\mathcal{R}}^{min}(x) \geq S_{\mathcal{R}}^{max}(y)$ holds for all $y \neq x$;
(2) x is a possible winner for $\mathcal{R}$ w.r.t. F_s if and only if $S_{\mathcal{R}}^{max}(x) \geq S_{\mathcal{R}}^{min}(y)$ holds for all $y \neq x$;

Proof: **1,$\Leftarrow$:** Suppose that x is not a necessary winner for $\mathcal{R}$ w.r.t. F_s. Then there exists an extension $\mathcal{T}$ of $\mathcal{R}$ and a $y \neq x$ such that $S(y,\mathcal{T}) > S(x,\mathcal{T})$, thus $S_{\mathcal{R}}^{max}(y) \geq S(y,\mathcal{T}) > S(x,\mathcal{T}) \geq S_{\mathcal{R}}^{min}(x)$, which contradicts the assumption that $S_{\mathcal{R}}^{min}(x) \geq S_{\mathcal{R}}^{max}(y)$ holds for all $y \neq x$.

1, $\Rightarrow$: Let x be a necessary winner. Then, there don't exists an $\mathcal{T} \in Ext(\mathcal{R})$ and there don't exists an $y \neq x$ such that $S(x,\mathcal{T}) < S(y,\mathcal{T})$. Since $S_{\mathcal{R}}^{min}(x) \leq S(x,\mathcal{T})$ and $S_{\mathcal{R}}^{max}(y) \geq S(y,\mathcal{T})$ we have that there don't exists an $y \neq x$ such that $S_{\mathcal{R}}^{min}(x) < S_{\mathcal{R}}^{max}(y)$.

2: analogously to 1. ■

$S_{\mathcal{R}}^{min}(x)$ considers the worst case and $S_{\mathcal{R}}^{max}(x)$ the best case for a scoring value for x. Hence, x is a necessary winner whenever the worst value is higher than the best value and x is a possible winner whenever the best value is higher than the worst value for x. Furthermore, we get the following result:

Corollary 1 *Possible and necessary winners for positional scoring procedures can be computed in polynomial time.*

For each partial order R_i and each candidate x, we just have to compute the number of candidates dominated by x and dominating x in R_i, which lead to the exact bound $\mathcal{O}(n{*}m^2)$.

Example 1 *Let us consider the following example where we have candidates $X = \{x_1, x_2, x_3, x_4\}$ and $p+q$ voters. The first group of p voters have the preferences $R_i = \{x_1 > x_2 > x_4, x_1 > x_3 > x_4\}$, for $0 \leq i \leq p$ and the other q voters have the preferences $R_i = \{x_3 > x_2 > x_1\}$, for $p+1 \leq i \leq p+q$. For the second group of voters, nothing is known about the position of x_4 with respect to other candidates (it is fully incomparable to them all). For the Borda procedure, we get*

	$S_{\mathcal{R}}^{min}$	$S_{\mathcal{R}}^{max}$
x_1	$3p$	$3p+q$
x_2	$p+q$	$2(p+q)$
x_3	$p+2q$	$2p+3q$
x_4	0	$3q$

Hence, there are no necessary winners. Candidate x_1 is a possible winner whenever $2p \geq q$, x_2 is possible if $2q \geq p$, x_3 is possible if $3q \geq p$, and x_4 is a possible winner if $q \geq p$.

For the plurality voting procedure we get

	$S_{\mathcal{R}}^{min}$	$S_{\mathcal{R}}^{max}$
x_1	p	p
x_2	0	0
x_3	0	q
x_4	0	q

Hence, we get: If $p > q$ then x_1 is the only possible winner and necessary winner. If $p = q$ then x_1, x_3 and x_4 are possible winners and x_1 is a necessary winner. If $p < q$ then x_3 and x_4 are possible winners and there is no necessary winner.

5 Condorcet winners

Recall that a candidate x is a *Condorcet winner* for a complete profile $\mathcal{T} = \langle \succeq_1, \ldots, \succeq_n \rangle$ if and only if for all $y \neq x$, $\#\{i \mid x \succ_i y\} > \frac{n}{2}$.[4]

Analogously to positional scoring procedures, we define upper and lower bounds for sets of Condorcet winners in case of partial preference profiles.

Definition 2 *Let $\mathcal{R}$ be an (incomplete) preference profile. Then,*

- *$x \in X$ is a* necessary Condorcet winner *for $\mathcal{R}$ if and only if for all $\mathcal{T} \in Ext(\mathcal{R})$, x is a Condorcet winner for $\mathcal{T}$.*
- *$x \in X$ is a* possible Condorcet winner *for $\mathcal{R}$ if and only if there exists a $\mathcal{T} \in Ext(\mathcal{R})$ such that x is a Condorcet winner for $\mathcal{T}$.*

Again, let us first focus on the worst and the best cases, this time by defining, for a *pair* of candidates (x,y), the number of voters for which x is preferred to y in the worst and in the best cases when considering all complete extensions of $\mathcal{T}$. If $\mathcal{T}$ is a collection of linear orders, let us first define

$$N_{\mathcal{T}}(x,y) = \#\{i \mid x >_i y\} - \#\{i \mid y >_i x\}$$

Then

$$N_{\mathcal{R}}^{min}(x,y) = \min_{\mathcal{T} \in Ext(\mathcal{R})} N_{\mathcal{T}}(x,y)$$
$$N_{\mathcal{R}}^{max}(x,y) = \max_{\mathcal{T} \in Ext(\mathcal{R})} N_{\mathcal{T}}(x,y)$$

$N_{\mathcal{R}}^{min}(x,y)$ (resp. $N_{\mathcal{R}}^{max}(x,y)$) corresponds to the the worst (resp. best) case for x among extensions of $\mathcal{R}$. Again, these bounds can be computed in polynomial time as follows:

Proposition 3 *Let $\mathcal{R}$ be a (partial) preference profile and x, y two distinct candidates from X. We define*

$$N_{R_i}^{max}(x,y) = \begin{cases} +1 & \text{if not } (y \geq_i x); \\ -1 & \text{if } y >_i x \end{cases} \quad \text{and}$$
$$N_{R_i}^{min}(x,y) = \begin{cases} +1 & \text{if } x >_i y; \\ -1 & \text{if not } (x \geq_i y) \end{cases}$$

Then

1. *$N_{\mathcal{R}}^{min}(x,y) = \sum_{i=1}^{n} N_{R_i}^{min}(x,y)$ and $N_{\mathcal{R}}^{max}(x,y) = \sum_{i=1}^{n} N_{R_i}^{max}(x,y)$;*

[4] Or equivalently, $\#\{i \mid x \succ_i y\} > \#\{i \mid y \succ_i x\}$. In the more general case when indifferences are allowed, this equivalence no longer holds and the latter expression is chosen as the usual definition of a Condorcet winner.

2. *x is a necessary Condorcet winner for $\mathcal{R}$ if and only if* $\forall y \neq x,\ N^{min}_{\mathcal{R}}(x,y) > 0$.

3. *x is a possible Condorcet winner for $\mathcal{R}$ if and only if* $\forall y \neq x,\ N^{max}_{\mathcal{R}}(x,y) > 0$.

Proof: **1:** Let us first show that $\sum_{i=1}^{n} N^{min}_{R_i}(x,y) = \min_{\mathcal{T}\in Ext(\mathcal{R})} N_{\mathcal{T}}(x,y)$. We have

$$\begin{aligned}\textstyle\sum_{i=1}^{n} N^{min}_{R_i}(x,y) = \ & \#\{i \mid x >_i y\} \\ & -\#\{i \mid not(x \geq_i y)\}\end{aligned}$$

Let $\mathcal{T} = \langle T_1,\dots,T_n\rangle \in Ext(\mathcal{R})$. We have $N_{\mathcal{T}}(x,y) = \#\{i \mid x >_{T_i} y\} - \#\{i \mid y >_{T_i} x\}$. Now, $x >_i y$ implies $x >_{T_i} y$, therefore (a) $\#\{i \mid x >_{T_i} y\} \geq \#\{i \mid x >_i y\}$ Next, $y >_{T_i} x$ implies not($x \geq_i y$), therefore (b) $\#\{i \mid y >_{T_i} x\} \leq \#\{i \mid \text{not}(x \geq_i y)\}$. (a) and (b) give (c) $\#\{i \mid x >_{T_i} y\} - \#\{i \mid y >_{T_i} x)\} \geq \#\{i \mid x >_i y\} - \#\{i \mid not(x \geq_i y)\}$, which is equivalent to $N_{\mathcal{T}}(x,y) \geq \#\{i \mid x >_i y\} - \#\{i \mid not(x \geq_i y)\}$. Since this holds for all $\mathcal{T} \in Ext(\mathcal{R})$, we get $\min_{\mathcal{T}\in Ext(\mathcal{R})} N_{\mathcal{T}}(x,y) \geq \sum_{i=1}^{n} N^{min}_{R_i}(x,y)$.

To show the inequality on the reverse direction, we consider, as in the proof of Proposition 1, the worst-case (for x) complete extension of R_i: for each i, let $(R_i)^{\downarrow}_x = R_i \cup \{(z,x) \mid z \neq x, not\ (x >_i z)\}$. Let $\mathcal{R}^{\downarrow}_x = \langle (R_1)^{\downarrow}_x,\dots,(R_n)^{\downarrow}_x\rangle$. We have $N_{\mathcal{R}^{\downarrow}_x}(x,y) =$

$$\begin{aligned} & \#\{i \mid x >_{(R_i)^{\downarrow}_x} y\} - \#\{i \mid y >_{(R_i)^{\downarrow}_x} x)\} \\ = \ & \#\{i \mid x >_{(R_i)^{\downarrow}_x} y\} - \#\{i \mid \text{not}(x \geq_{(R_i)^{\downarrow}_x} y)\} \\ = \ & \#\{i \mid x >_{R_i} y\} - \#\{i \mid \text{not}(x \geq_{R_i} y)\} \\ = \ & \textstyle\sum_{i=1}^{n} N^{min}_{R_i}(x,y)\end{aligned}$$

Therefore, $\min_{\mathcal{T}\in Ext(\mathcal{R})} N_{\mathcal{T}}(x,y) \leq \sum_{i=1}^{n} N^{min}_{R_i}(x,y)$.

The proof for $N^{max}_{\mathcal{R}}(x,y) = \sum_{i=1}^{n} N^{max}_{R_i}(x,y)$ is similar.

2, ⇐: Assume that x is not a necessary winner. Then, there exists an $\mathcal{T} \in Ext(\mathcal{R})$ and an $y \neq x$ such that $\#\{i \mid x >_{T_i} y\} \leq \#\{i \mid y >_{T_i} x\}$. Hence, $N_{\mathcal{T}}(x,y) \leq 0$. That implies that $\min_{\mathcal{T}\in Ext(\mathcal{R})} N_{\mathcal{T}}(x,y) \leq 0$ and thus, $N^{min}_{\mathcal{R}}(x,y) \leq 0$ which is a contradiction to the assumption that $N^{min}_{\mathcal{R}}(x,y) > 0$ holds. Hence, x is a necessary Condorcet winner.

2, ⇒: Let x be a necessary Condorcet winner. Assume that there exists an $y \neq x$ such that $N^{min}_{\mathcal{R}}(x,y) \leq 0$ holds for some x. That is, we have $\#\{i \mid x >_i y\} \leq \#\{i \mid not(x \geq_i y)\}$. Hence, there exists an $\mathcal{T} \in Ext(\mathcal{R})$ such that $N_{\mathcal{T}}(x,y) \leq 0$. Hence, x is no Condorcet winner, which is a contradiction to the assumption.

3: similar to the proof for necessary winners. ■

If x is strictly preferred to y, then $N^{min}_{\mathcal{R}}(x,y)$ and $N^{max}_{\mathcal{R}}(x,y)$ assign the value 1 as in the case for complete preferences. Furthermore, if candidates x and y are incomparable, the function $N^{max}_{R_i}(x,y)$ assigns the value 1 and $N^{min}_{R_i}(x,y)$ the value -1. This follows the intuition that $N^{max}_{\mathcal{R}}$ covers the "best" case and $N^{min}_{\mathcal{R}}$ the "worst" case for candidates x and y. Hence, if $N^{min}_{\mathcal{R}}(x,y) > 0$ for all $y \neq x$, then in the worst case, strictly more voters prefer x strictly over y than y over x. In this case, x is a necessary Condorcet winner. Whenever $N^{max}_{\mathcal{R}}(x,y) > 0$, there exists a complete extension $\mathcal{T}$ of $\mathcal{R}$ such that x is a Condorcet winner and hence, x is a possible Condorcet winner.

Example 2 *Let us reconsider Example 1.*

We get that x_1 is a necessary Condorcet winner if $p > q$. The other candidates become never a necessary Condorcet winner. Furthermore, x_1 is a possible Condorcet winner if $p > q$ and x_3 and x_4 are possible Condorcet winners if $q > p$. Candidate x_2 is not a possible Condorcet winner.

As stated in Section 4, we can compute necessary and possible Condorcet winners in polynomial time.

Corollary 2 *Possible and necessary Condorcet winners can be computed in polynomial time.*

One may wonder whether this way of determining possible and necessary winners just by computing lower and upper bounds of scores, which works for scoring procedures and Condorcet winners, extends to Condorcet-consistent voting procedures such as the Simpson or the Copeland procedures [Brams and Fishburn, 2003]. Unfortunately, this is not so simple, as the method consisting in computing lower and upper bounds does not suffice[5]. Computing possible and necessary winners for such procedures might be NP-hard and coNP-hard. This issue is left for further research.

6 Related issues

We now investigate the links between possible and necessary winners and some issues such as vote elicitation and manipulation.

6.1 Manipulation

The Gibbard-Satterthwaite theorem states that any vote procedure can be manipulated, or in other terms, that voters sometimes have an interest to report unsincere preferences. Manipulation was recently revisited from the computational point of view [Conitzer and Sandholm, 2002a; Conitzer *et al.*, 2003; Conitzer and Sandholm, 2003]: given (a) a coalition of voters $J \subseteq I$; (b) a candidate $x \in X$ and (c) the individual profiles $\mathcal{R}_{I\setminus J} = \langle R_j\rangle_{j\in I\setminus J}$ of the voters in $I \setminus J$:

- A *constructive manipulation* for x by J given $\mathcal{R}_{I\setminus J}$ (with respect to a given vote procedure F) is a way for the voters in J to cast their votes such that x is guaranteed to win the election, that is, a set of individual profiles $\mathcal{R}_J$ such that $F(\langle \mathcal{R}_{J\setminus I}, \mathcal{R}_J\rangle) = \{x\}$.
- A *destructive manipulation* for x by J given $\mathcal{R}_{I\setminus J}$ (with respect to a given vote procedure F) is a way for the voters in J to cast their votes such that x is guaranteed not to win the election, that is, a set of individual profiles $\mathcal{R}_J$ such that $x \notin F(\langle \mathcal{R}_{J\setminus I}, \mathcal{R}_J\rangle)$.

[5] Consider for instance the Simpson (or *maximin*) procedure, consisting of choosing the candidates maximizing the Simpson score $S_{\mathcal{T}}(x) = \min_{y\neq x} N_{\mathcal{T}}(x,y)$. Then, given a partial preference profile $\mathcal{R}$, we may compute in polynomial time a lower bound $S^{min}_{\mathcal{R}}(x) = \min_{\mathcal{T}\in Ext(\mathcal{R})} S_{\mathcal{T}}(x)$ and an upper bound $S^{max}_{\mathcal{R}}(x) = \max_{\mathcal{T}\in Ext(\mathcal{R})} S_{\mathcal{T}}(x)$. However, even if, for instance, $S^{min}_{\mathcal{R}}(x) > S^{min}_{\mathcal{R}}(y)$ for all y implies that x is a necessary winner, the converse implication is not guaranteed to hold, because it may be the case that no extension of $\mathcal{R}$ simultaneously gives a minimal score to x and a maximal score to y.

We then have the following easy results (the proof of which are omitted). Let $R_\emptyset = \{(x, x) \mid x \in X\}$.

Proposition 4 *Let F be a voting procedure, $J \subseteq I$ be a coalition of voters, $x \in X$ and $\mathcal{R}_{I \setminus J} = \langle R_j \rangle_{j \in I \setminus J}$. We let $\mathcal{R}^* = \langle R_i^* \rangle_{i \in I}$ where $R_i^* = R_i$ if $i \in I \setminus J$ and $R_i^* = R_\emptyset$ if $i \in J$.*

1. *there is a constructive manipulation for x by J given $\mathcal{R}_{I \setminus J}$ if and only if $PW_F(\mathcal{R}^*) = \{x\}$;*
2. *there is a destructive manipulation for x by J given $\mathcal{R}_{I \setminus J}$ if and only if $x \notin NW_F(\mathcal{R}^*)$.*

Thus, deciding whether there is a constructive or a destructive manipulation for a given candidate is a subproblem of voting with partial preference relations. As an obvious corollary, whenever computing necessary and possible winners is polynomial, then deciding whether there is a (constructive/destructive) manipulation is polynomial as well[6] .

6.2 Elicitation

Given a set of individual profiles $\mathcal{R}_J = \langle R_j \rangle_{j \in J}$ corresponding to a subset of voters $J \subseteq I$ who have already expressed their votes. *Vote elicitation* [Conitzer and Sandholm, 2002b] consists in determining, whether (a) the outcome of the vote can be determined without needing any further information and (b) which information must be asked to which voter. We generalize these notions to the more general situation where the initial knowledge about the votes is any partial preference profile: given a partial preference profile $\mathcal{R}$, the elicitation task is over iff it is useless to learn more about the voter's preferences, that is, the outcome of the vote will be the same in any complete extension of $\mathcal{R}$: for any $\mathcal{T}, \mathcal{T}' \in Ext(\mathcal{R})$, $F(\mathcal{T}) = F(\mathcal{T}')$. This condition is easily shown to be equivalent to the fact that possible and necessary winners coincide:

Proposition 5 *Given a voting procedure F and a partial preference profile $\mathcal{R}$, the elicitation process is over if and only if $PW_F(\mathcal{R}) = NW_F(\mathcal{R})$.*

7 Conclusion

In this paper we made first steps towards computing the outcome of voting procedures when the voters' preferences are incomplete, and we pointed connections to vote manipulation and elicitation.

For the sake of simplicity, we required the voters' preferences to be antisymmetric. However, definitions of possible and necessary winners carry on to the more general case where voters' incomplete preferences are weak orders (allowing for indifferences), provided that the voting procedure F allows for indifferences as well. Especially, possible and necessary Condorcet winners can still be defined, and computed in polynomial time.

Further work obviously includes the investigation of other voting rules, as briefly evoked at the end of Section 5. Another interesting issue would consist in defining a middle way between possible and necessary winners, by counting the number of extensions in which a candidate is a winner. This probabilistic criterion will probably be much harder to compute than the extremely optimistic and pessimistic criteria underlying the notions of possible and necessary winners.

Incompleteness here refers only to *preferences*. Another place where incompleteness may be relevant is *in the voting procedure itself*: this is the way followed by [Conitzer and Sandholm, 2003], who introduce some uncertainty in the way the voting procedure will be applied so as to make manipulation more difficult. Although both issues are significantly different, it is worth considering whether studying both in a unifying framework would be relevant.

Acknowledgements The first author was partially supported by the German Science Foundation (DFG) under grant SCHA 550/6-4, TP C and by the EC under project IST-2001-37004 WASP.

References

[Boutilier *et al.*, 2004] C. Boutilier, R. I. Brafman, C. Domshlak, H. Hoos, and D. Poole. Preference-based constrained optimization with CP-nets. *Computational Intelligence*, 20(2):137–157, 2004.

[Boutilier, 1994] C. Boutilier. A logical approach to qualitative decision theory. In *Proceedings of UAI-94*, 1994.

[Brams and Fishburn, 2003] S. Brams and P. Fishburn. *Handbook of Social Choice and Welfare*, volume 1, chapter Voting procedures. Elsevier, 2003.

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

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

[Conitzer and Sandholm, 2003] V. Conitzer and T. Sandholm. Universal voting protocols tweaks to make manipulation hard. In *Proceedings of IJCAI-03*, pages 781–788, 2003.

[Conitzer *et al.*, 2003] V. Conitzer, J. Lang, and T. Sandholm. How many candidates are required to make an election hard to manipulate? In *Proceedings of TARK-03*, pages 201–214, 2003.

[Davenport and Kalagnanam, 2004] A. Davenport and J. Kalagnanam. A computational study of the Kemeny rule for preference aggregation. In *Proceedings of AAAI-04*, 2004.

[Hemaspaandra *et al.*, 2004] E. Hemaspaandra, H. Spakowski, and J. Vogel. The complexity of Kemeny elections. Technical report, 2004.

[Lang, 2004] J. Lang. Logical preference representation and combinatorial vote. *Annals of Mathematics and Artificial Intelligence*, 42(1):37–71, 2004.

[Rossi *et al.*, 2004] F. Rossi, K. Venable, and T. Walsh. mCP-nets: representing and reasoning with preference of multiple agents. In *Proceedings of AAAI-04*, 2004.

[Rothe *et al.*, 2003] J. Rothe, H. Spakowski, and J. Vogel. Exact complexity of the winner determination problem in young elections. *Theory of Computing Systems*, 36:375–386, 2003.

[Tan and Pearl, 1994] S.-W. Tan and J. Pearl. Qualitative decision theory: specification and representation of preferences under uncertainty. In *Proceedings of KR-94*, 1994.

[6]Note that the NP-hardness results of [Conitzer and Sandholm, 2002a] do not apply here, since they apply to *weighted* votes.