Preference Aggregation over Restricted Ballot Languages: Sincerity and Strategy-Proofness

Ulle Endriss
University of Amsterdam

Maria Silvia Pini
University of Padova

Francesca Rossi
University of Padova

K. Brent Venable
University of Padova

Abstract

Voting theory can provide useful insights for multiagent preference aggregation. However, the standard setting assumes voters with preferences that are total orders, as well as a ballot language that coincides with the preference language. In typical AI scenarios, these assumptions do not hold: certain alternatives may be incomparable for some agents, and others may have their preferences encoded in a format that is different from how the preference aggregation mechanism wants them. We study the consequences of dropping these assumptions. In particular, we investigate the consequences for the important notion of strategy-proofness. While strategy-proofness cannot be guaranteed in the classical setting, we are able to show that there are situations in our more general framework where this is possible. We also consider computational aspects of the problem.

1 Introduction

Many AI scenarios involve dealing with several agents that declare their preferences over a set of possible decisions, and choosing the decision that satisfies such preferences in the best way [Wooldridge, 2002; Walsh, 2007]. Examples can be found in several domains, such as recommender systems, configuration, combinatorial auctions, distributed planning or scheduling, automated assistants, catalogue search, timetabling, and assistive technology.

When considering the preferences and deciding which decision to take, the setting is in principle similar to what voting theory considers: several voters express, via a ballot, their preferences over a set of candidates, and a candidate is chosen in one of many possible ways (via a voting rule) to decide who the winning candidate is. Thus it is natural to explore the application of classical and more recent results in voting theory to such AI settings. The following are three classical voting procedures [Brams and Fishburn, 2002]:

- *Plurality:* Each voter votes for exactly one candidate and the candidate receiving the most votes wins.

- *Borda:* Each voter submits a complete ranking of all m candidates. If a candidate is ranked highest by some

voter, he receives $m-1$ points; if he is ranked second, he receives $m-2$ points, and so on. The candidate receiving the most points wins the election.

- *Approval:* Each voter approves of as many candidates as he wishes by submitting a set of candidate names. The candidate receiving the most approvals wins.

While originally defined for political and social use, such procedures can also be used in scenarios where autonomous software agents express their preferences.

Many results in voting theory rest, among other things, on two important assumptions: First, voters have *preferences* that are *total orders* over the set of candidates. That is, they are expected to be able to strictly rank all candidates. Second, voters are supposed to vote by reporting their preferences to the election chair (and they may or may not do so truthfully). That is, the *ballots* used to communicate a voter's input to the mechanism entrusted with computing the winner(s) of an election are the same kind of structure as those used to represent actual preferences (both are total orders).

As an example, consider the well-known Gibbard-Satterthwaite Theorem [Gibbard, 1973; Satterthwaite, 1975]. It states that there can be no voting rule for three or more candidates that is both *non-dictatorial* (i.e., that does not always elect the preferred candidate of some distinguished voter, the dictator) and *strategy-proof* (i.e., that never gives a voter an incentive to manipulate by misrepresenting their preferences). Here a voting rule is taken to be a function mapping any given profile of total orders (the ballots) to a winning candidate. Whether or not a voter has an incentive to manipulate is evaluated with respect to his actual preference, which is also assumed to be a total order.

In typical AI scenarios, above assumptions often do not hold: some alternatives may be *incomparable* [Pini *et al.*, 2007], and agents may have their preferences encoded in a way that is not how the preference aggregation mechanism wants them. For example, agents may model their preferences as preorders (thus with possible indifferences and/or incomparabilities), while the ballot language may only allow for total orders, as in the Borda rule. Even if agents use total orders, they may wish to declare only a portion of it, due to, say, privacy reasons or elicitation costs, or they may be forced to give a more restrictive structure, as in approval voting.

Hence, a voter may sometimes be *unable* to vote truthfully

in the classical sense, as reporting their true preferences may simply not be allowed given the ballot language in place. If reporting one's true preference is *impossible*, the definition of strategy-proofness should be relaxed so as not to label as manipulators those voters that submit a ballot that is admitted by the language and that is "as close as possible" to their true preference. To formalise this idea we shall give a (in fact several) definition(s) of what constitutes a *sincere* ballot in our generic framework and we then define strategy-proofness in terms of sincerity rather than truthfulness. In the special case where expressing one's true preferences in the ballot language *is* possible (as in the classical setting, among others), our notion of sincerity will reduce to simple truthfulness and we arrive at the standard definition of strategy-proofness.

While in the classical setting strategy-proofness is unachievable (for non-dictatorial rules), we will show that this is not the case in our more general setting: it is possible, in some cases, to define and use preference aggregation mechanisms that never give an agent an incentive to vote insincerely.

The remainder of this paper is structured as follows. Section 2 introduces the general framework and basic notation, and Section 3 shows how to define a voting procedure in this framework. Section 4 defines three notions of sincerity and establishes relationships between them. Section 5 introduces our generalised definition of strategy-proofness and proves positive results for two types of voting scenarios. Section 6 discusses computational aspects of these results. (For lack of space, some proofs are only sketched.)

2 Orders, Preferences, and Ballots

Throughout this paper, $\mathcal{P}$ denotes the set of all possible *preference orders* and $\mathcal{B}$ the set of all *valid ballots* (or the *ballot language*). The classical setting is obtained when $\mathcal{P}$ is the set of all total orders and $\mathcal{B} = \mathcal{P}$. Moreover, let $\mathcal{C}$ be the set of *candidates* and $m = |\mathcal{C}|$ the number of candidates. Further, let $\mathcal{N}$ be the set of *voters* and $n = |\mathcal{N}|$ the number of voters.

2.1 Orders

We use standard terminology to talk about different types of orders [Roberts, 1979]. A *preorder* is a binary relation that is reflexive and transitive. A *partial order* is a preorder that is antisymmetric. A *weak order* is a preorder that is complete. A *total order* is a partial order that is complete. That is, a preorder will either strictly rank a pair of elements, declare them indifferent, or consider them incomparable. Partial orders exclude the case of indifference, while weak orders do not allow for incomparabilities. Finally, a total order will strictly rank any two (distinct) elements. The set of all preorders includes all of the other aforementioned classes of relations.

Any preorder p over $\mathcal{C}$ induces a partition of $\mathcal{C}^2$ into four sets $R_p^{\prec}$ ("strictly worse"), $R_p^{\succ}$ ("strictly better"), $R_p^{\sim}$ ("indifferent"), and $R_p^{\bowtie}$ ("incomparable"), defined as follows:

(1) $R_p^{\prec} = \{(x,y) \in \mathcal{C}^2 \mid (x,y) \in p \text{ and } (y,x) \notin p\}$

(2) $R_p^{\succ} = \{(x,y) \in \mathcal{C}^2 \mid (x,y) \notin p \text{ and } (y,x) \in p\}$

(3) $R_p^{\sim} = \{(x,y) \in \mathcal{C}^2 \mid (x,y) \in p \text{ and } (y,x) \in p\}$

(4) $R_p^{\bowtie} = \{(x,y) \in \mathcal{C}^2 \mid (x,y) \notin p \text{ and } (y,x) \notin p\}$

2.2 Preferences

The set $\mathcal{P}$ of preferences that are considered possible as true preferences for the voters will always be a non-empty set of preorders. This is the only general constraint we impose.

Of special interest is the case when $\mathcal{P}$ is the set of all total orders, as in the classical setting. Another special setting that we are going to analyse is what we call "*2-level preorders*": p is a 2-level preorder if it divides $\mathcal{C}$ into (at most) two levels, and each candidate in the top level is preferred to each candidate in the bottom level. Between candidates at the same level we may have either indifference or incomparability.

2.3 Ballot Languages

The set $\mathcal{B}$ of valid ballots also has to be a non-empty set of preorders. In addition, we require $\mathcal{B}$ to satisfy the following condition: for all total orders p over $\mathcal{C}$ there must exists a $b \in \mathcal{B}$ such that $(x,y) \in R_p^{\prec}$ implies $(x,y) \notin R_b^{\succ}$ for all $x,y \in \mathcal{C}$. This condition is satisfied, for instance, if $\mathcal{B}$ includes a ballot that declares any two candidates incomparable (or indifferent). The reason for this condition is that we do not want to force a particular strict ordering between any two candidates *a priori*. However, we do allow ballot languages to enforce indifferences or incomparabilities; and we also allow them to enforce *some* strict ordering (only the direction of this ordering must always be left to the voter). We will sometimes refer to the following special ballot languages:

(1) *Borda ballots:* A Borda ballot b is a total order over $\mathcal{C}$. The name derives from the fact that this kind of ballot is what is needed to apply the Borda rule.

(2) *Plurality ballots:* A plurality ballot b is a partial order over $\mathcal{C}$ with exactly one undominated candidate dominating all other candidates, which in turn are mutually incomparable. That is, a plurality ballot has the form of a tree with only leaf nodes and a root. This kind of ballot best reflects the information requirements for plurality voting: a voter has to single out one top candidate and does not make any further value judgements.

(3) *Approval ballots:* An approval ballot b is a 2-level partial order over $\mathcal{C}$ (so the candidates at the same level are mutually incomparable). This directly corresponds to the standard form of balloting in approval voting: voters can distinguish approved and non-approved voters, but do not make any value judgements beyond that. We call b an *abstention ballot* if it has only a single level (i.e., if all candidates are undominated).

We call a ballot language *neutral* if permuting the candidate names on a valid ballot will never render that ballot invalid.

3 A Generic Voting Procedure

Our aim now is to define a generic voting rule over preorders that reduces to the well-known rules of plurality, Borda, and approval in case the ballot language is restricted to the corresponding types of ballots. The first method that comes to mind is to give each candidate as many points as they dominate other candidates in the ballots submitted. This method works for Borda and plurality, but not for approval voting.

The generic voting rule defined next solves this problem. It computes the score of voter i awarded to candidate c as the length of the longest path from c down to some candidate at the bottom of the ballot of i. The candidate with the highest sum of such scores wins.

Definition 1 (Path below candidate) *Given a ballot b, a path below candidate $c \in \mathcal{C}$ is a set $C \subset \mathcal{C}$ with some $c^{\perp} \in C$ such that $(c, c') \in R_b^{\succ}$ and $(c', c^{\perp}) \in R_b^{\succ} \cup R_b^{\sim}$ for all $c' \in C$. The length of path C below c is $|C|$.*

Definition 2 (Longest-path voting) *Given a ballot $b \in \mathcal{B}$, the score of a candidate $c \in \mathcal{C}$ with respect to b is the length of the longest path below c. For a set of ballots, the score of c is the sum of the scores with respect to the ballots in the set. The longest-path voting procedure declares the candidates with maximal longest-path score the winners.*

It is not hard to see that this satisfies our requirements:

Theorem 1 *The longest-path voting procedure reduces to the Borda rule if applied to Borda ballots, to the plurality rule if applied to plurality ballots, and to the approval voting rule if applied to approval ballots.*

In this paper we will focus on Borda, approval, and plurality, which are instances of the longest-path voting procedure just defined, even if the idea of voting with ballots that are preorders can also be extended to other voting rules, such as the family of *Condorcet-consistent rules* [Brams and Fishburn, 2002], including the Copeland rule and the Dodgson rule.

4 Three Notions of Sincerity

In some cases we have clear intuitions what constitutes a *sincere* vote. Certainly, when $\mathcal{P} = \mathcal{B}$, as in most work in classical voting theory, we would consider a ballot sincere only if it is equal to the voter's true preference. However, when we relax the assumption that $\mathcal{P} = \mathcal{B}$, different notions of sincerity may be reasonable. For approval voting, the standard definition of sincerity says that a ballot (a set of approved candidate names) is sincere with respect to a given true preference p if p ranks each approved candidate at least as high as any of the non-approved candidates [Brams and Fishburn, 2002].

Here we propose three definitions of sincerity. The first of these only imposes minimal constraints on a sincere ballot. All we require is that there is no explicit contradiction as far as pairs of strictly ordered candidates are concerned.

Definition 3 (Minimal sincerity) *Let p be a preorder and $\mathcal{B}$ a ballot language. A ballot $b \in \mathcal{B}$ is minimally sincere wrt. p if $R_b^{\prec} \cap R_p^{\succ} = \emptyset$. We write $b \in \text{SIN}_{\mathcal{B}}^{\min}(p)$.*
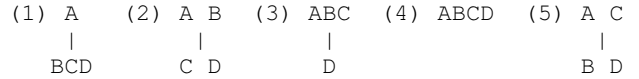
Our second definition furthermore requires that the ballot and the true preference agree as much as possible, where "as much as possible" is interpreted as maximality with respect to *set-inclusion*. For the next two definitions, let $\text{AGR}_p(b)$ be $(R_b^{\prec} \cap R_p^{\prec}) \cup (R_b^{\succ} \cap R_p^{\succ}) \cup (R_b^{\sim} \cap R_p^{\sim}) \cup (R_b^{\bowtie} \cap R_p^{\bowtie})$.

Definition 4 (Qualitative sincerity) *Let p be a preorder and $\mathcal{B}$ a ballot language. A ballot $b \in \mathcal{B}$ is qualitatively sincere wrt. p if $b \in \text{SIN}_{\mathcal{B}}^{\min}(p)$ and there is no $b' \in \text{SIN}_{\mathcal{B}}^{\min}(p)$ such that $\text{AGR}_p(b) \subset \text{AGR}_p(b')$. We write $b \in \text{SIN}_{\mathcal{B}}^{\text{qual}}(p)$.*

Our third proposal is similar, but now we require maximality with respect to the *number* of agreements.

Definition 5 (Quantitative sincerity) *Let p be a preorder and $\mathcal{B}$ a ballot language. A ballot $b \in \mathcal{B}$ is quantitatively sincere wrt. p if $b \in \text{SIN}_{\mathcal{B}}^{\min}(p)$ and there is no $b' \in \text{SIN}_{\mathcal{B}}^{\min}(p)$ such that $|\text{AGR}_p(b)| < |\text{AGR}_p(b')|$. We write $b \in \text{SIN}_{\mathcal{B}}^{\text{quan}}(p)$.*

Let us now give an example that shows that each of our three definitions can yield a different set of sincere ballots. Suppose there are four candidates and our voter's true preferences are $A \succ B \succ C \succ D$. If we restrict attention to approval ballots, there are 15 syntactically valid ballots (our voter may approve of any non-empty subset of the set of four candidates). Five of these are depicted below:

```
(1)  A       (2)  A  B    (3)  ABC    (4)  ABCD    (5)  A  C
      |             |            |                        |
     BCD           C  D          D                       B  D
```

Candidates shown on the same level are understood to be incomparable. Ballots (1)–(4) are the four sincere ballots according to the standard definition of sincerity in the approval voting literature. These four ballots are also minimally sincere, while ballot (5) is not. (1)–(3) are furthermore qualitatively sincere: for none of them we would be able to add a further strict comparability relation without violating the ordering of the true preference. Ballot (4), which corresponds to the voter abstaining, is not qualitatively sincere, because this ballot could be further refined, e.g., by moving to (3) instead. Ballot (2) is the only quantitatively sincere ballot. It includes four of the strictly ordered pairs of the true preference, while (1) and (3) only include three strictly ordered pairs each.

Next, we establish some basic properties of our three notions of sincerity. We first show that there is a natural order over these notions in terms of how restrictive they are:

Theorem 2 *Let p be a preorder and let $\mathcal{B}$ be a ballot language. Then $\text{SIN}_{\mathcal{B}}^{\min}(p) \supseteq \text{SIN}_{\mathcal{B}}^{\text{qual}}(p) \supseteq \text{SIN}_{\mathcal{B}}^{\text{quan}}(p) \supset \emptyset$.*

Proof. (1) $\text{SIN}_{\mathcal{B}}^{\min}(p) \supseteq \text{SIN}_{\mathcal{B}}^{\text{qual}}(p)$ is immediate from the definition of qualitative sincerity. (2) $\text{SIN}_{\mathcal{B}}^{\text{qual}}(p) \supseteq \text{SIN}_{\mathcal{B}}^{\text{quan}}(p)$ follows from the fact that any set that is maximal with respect to cardinality must certainly be maximal with respect to set-inclusion. (3) $\text{SIN}_{\mathcal{B}}^{\text{quan}}(p) \supset \emptyset$ is a direct consequence of our constraint on $\mathcal{B}$ requiring that a ballot language may not be so restrictive so as to force one particular strict ordering on a pair of candidates. Hence, there must always be at least one ballot b that does not violate $R_b^{\prec} \cap R_p^{\succ} = \emptyset$. □

Whenever the ballot language can express a voter's true preference, then that will be the only sincere ballot according to both the qualitative and the quantitative definition:

Theorem 3 *If $\mathcal{B} \supseteq \mathcal{P}$, then $\text{SIN}_{\mathcal{B}}^{\text{qual}}(p) = \text{SIN}_{\mathcal{B}}^{\text{quan}}(p) = \{p\}$ for all $p \in \mathcal{P}$.*

Proof. Clearly, $\text{AGR}_p(b) \subseteq \text{AGR}_p(p)$ for any ballot b. Hence, no b different from p could possibly be maximal with respect to set-inclusion and p must be the only element of $\text{SIN}_{\mathcal{B}}^{\text{qual}}(p)$. Equality with $\text{SIN}_{\mathcal{B}}^{\text{quan}}(p)$ then follows from Theorem 2. □

When the range of allowed ballots are the total orders, then all three notions of sincerity coincide:

Theorem 4 *If $\mathcal{B}$ is the set of all total orders, then we have* $\mathrm{SIN}_{\mathcal{B}}^{\min}(p) = \mathrm{SIN}_{\mathcal{B}}^{\mathrm{qual}}(p) = \mathrm{SIN}_{\mathcal{B}}^{\mathrm{quan}}(p)$ *for all preorders p.*

Proof. We only need to show $\mathrm{SIN}_{\mathcal{B}}^{\min}(p) \subseteq \mathrm{SIN}_{\mathcal{B}}^{\mathrm{quan}}(p)$; the rest then follows from Theorem 2. Let $b \in \mathrm{SIN}_{\mathcal{B}}^{\min}(p)$. As b is a total order and minimally sincere, we must have $R_b^{\prec} \supseteq R_p^{\prec}$ and $R_b^{\succ} \supseteq R_p^{\succ}$. Because b has to strictly rank all pairs of candidates, we also get $R_b^{\sim} = R_b^{\bowtie} = \emptyset$. It follows that $\mathrm{AGR}_p(b) = R_p^{\prec} \cup R_p^{\succ}$. Hence, no other ballot can achieve a higher number of agreements, and b must be quantitatively sincere, i.e., $b \in \mathrm{SIN}_{\mathcal{B}}^{\mathrm{quan}}(p)$. $\qquad\square$

Additional properties hold when we consider approval ballots, which are partial orders with at most two levels:

Theorem 5 *Let $\mathcal{P}$ be the set of total orders and let $\mathcal{B}$ be the set of approval ballots. Then the following hold:*

(a) *Minimal sincerity coincides with the standard notion of sincerity in approval voting.*

(b) *For all $p \in \mathcal{P}$, $\mathrm{SIN}_{\mathcal{B}}^{\mathrm{qual}}(p) = \mathrm{SIN}_{\mathcal{B}}^{\min}(p) \setminus \{\mathrm{ABS}\}$, where* ABS *is the abstention ballot in approval voting.*

(c) *For all $p \in \mathcal{P}$, the number of quantitatively sincere ballots $|\mathrm{SIN}_{\mathcal{B}}^{\mathrm{quan}}(p)|$ is 2 if m is odd (approving of the $\frac{m-1}{2}$ and $\frac{m+1}{2}$ top candidates, respectively) and 1 if m is even (approving of exactly the top half of $\mathcal{C}$).*

The proof is omitted for lack of space, but the example given earlier illustrates the intuitions.

5 Generalised Strategy-Proofness

As argued in the introduction, in the context of elections where voting by means of one's true preference may be impossible due to restrictions to the ballot language, a generalised definition of strategy-proofness is required.

A *voting correspondence* $f : \mathcal{B}^n \to 2^{\mathcal{C}}$ is a function mapping a ballot profile to a (non-empty) set of winning candidates (so far we have used the less technical term "voting procedure"). Roughly speaking, the standard definition of strategy-proofness says that a voting correspondence is strategy-proof if it never gives any voter an incentive to vote differently from reporting their *true preference* to the election chair. In our more general model, where we distinguish preferences from ballots, this ("reporting their true preference") is not a meaningful notion. Instead, we define a generalised notion of strategy-proofness that requires that a voter will never have an incentive to not vote by means of a *sincere ballot*.

To be able to make "having an incentive" precise a notion we need to be able to talk about a voter i who prefers one *set of winners* over another. This is not possible by means of that voter's preference $p \in \mathcal{P}$ alone, because p is defined over $\mathcal{C}$ rather than over $2^{\mathcal{C}}$. That is, we need to lift i's preferences over individual candidates to *preferences over sets of candidates*. There are different ways of doing this [Barberà *et al.*, 2004]. Denote by $\lhd_p$ the result of lifting p to a relation declared over non-empty sets of candidates. This will be one of the parameters of our definition of strategy-proofness. A widely followed approach is to define $\lhd_p$ as the smallest partial order over $2^{\mathcal{C}} \setminus \{\emptyset\}$ that satisfies the so-called *Gärdenfors axioms* [Barberà *et al.*, 2004]:

- $S \cup \{x\} \lhd_p S$ whenever $(x, y) \in R_p^{\prec}$ for all $y \in S$
- $S \lhd_p S \cup \{y\}$ whenever $(x, y) \in R_p^{\prec}$ for all $x \in S$

(Here, S is any non-empty set of candidates, and x and y represent candidates.) That is, under the Gärdenfors lifting, we say that a voter with a given preference relation over individual candidates will prefer the set of candidates Y over the set X (denoted $X \lhd_p Y$) if and only if Y can be obtained from X by a sequence of operations consisting of either removing the (strictly) least preferred candidate (except when only one candidate is left in the set) or adding a (strictly) more preferred candidate. Any other pairs of sets of candidates are considered incomparable. For example, if I like A more than B more than C, then $\{A, C\}$ is worse than $\{A\}$ (first axiom), while $\{A, C\}$ and $\{B\}$ are incomparable. Note that while the specific results given later on assume the Gärdenfors axioms, our definition of generalised strategy-proofness can be instantiated to any other notion of lifting as well.

A second parameter in our definition of strategy-proofness is the notion of sincerity used. We state the definition in its full generality and then discuss various instantiations for these parameters. In the next definition, let $\mathcal{N}$ be a set of voters, $\mathcal{C}$ a set of candidates, $\mathcal{P}$ a set of preference orders over $\mathcal{C}$, $\mathcal{B}$ a set of ballots over $\mathcal{C}$, $\mathrm{SIN}_{\mathcal{B}} : \mathcal{P} \to 2^{\mathcal{B}}$ a notion of sincerity, and $\lhd_{p_i}$ a lifting of preferences $p_i \in \mathcal{P}$ to preferences over sets of candidates.

Definition 6 (Generalised strategy-proofness) *A voting correspondence $f : \mathcal{B}^n \to 2^{\mathcal{C}}$ is g-strategy-proof if, for all voters $i \in \mathcal{N}$ with true preference $p_i \in \mathcal{P}$ and for all ballot vectors $b \in \mathcal{B}^n$, there exists a sincere ballot $b_i' \in \mathrm{SIN}_{\mathcal{B}}(p_i)$ such that $f(b_{-i}, b_i') \lhd_{p_i} f(b)$ does not hold.*[1]

In other words, f is g-strategy-proof if no voter will ever do worse by voting sincerely rather than insincerely—at least for one of the sincere ballots available to him. As $\lhd_{p_i}$ need not be a total order, requiring $f(b_{-i}, b_i') \lhd_{p_i} f(b)$ *not* to hold is not the same as requiring $f(b_{-i}, b_i') \unrhd_{p_i} f(b)$. Arguably, either option could be used to define strategy-proofness. The former means that voters may manipulate only if they have positive reason to do so. The latter means that voters may manipulate unless they have positive reason for not doing so. While Definition 6 uses the former option, the latter is the (implicit) choice made in [Endriss, 2007], discussing sincerity and manipulation in approval voting. Our choice is appropriate when incomparabilities in a voter's (lifted) preference are interpreted as the inability to compare alternatives; the other (more "pessimistic") approach is appropriate when incomparabilities are interpreted as uncertainty on behalf of the mechanism designer. As a consequence of this difference, the results on strategy-proofness reported below differ (are more positive) than those given in [Endriss, 2007].

Whether or not a particular voting procedure will be found to be g-strategy-proof will depend on several factors, besides the definition of the procedure itself. In particular, g-strategy-proofness will be more likely to hold when we use

- a smaller set of potential preferences $\mathcal{P}$; or
- a smaller set of valid ballots $\mathcal{B}$; or

---

[1] (b_{-i}, b_i') is the vector we obtain when we replace b_i in b by b_i'.

- a less restrictive notion of sincerity; or

- a definition of lifted preferences $\lhd_{p_i}$ that leaves more pairs of sets of candidates incomparable.

For example, if f is a voting *rule* (producing a single winner, i.e., the only reasonable definition of $\lhd_{p_i}$ is to identify it with p_i), $\mathcal{P}$ is the set of all total orders, $\mathcal{B} = \mathcal{P}$, and $\text{SIN}_{\mathcal{B}}(p) = \{p\}$, then Definition 6 reduces to the standard definition of strategy-proofness usually given in the context of the Gibbard-Satterthwaite Theorem [Gibbard, 1973; Satterthwaite, 1975]. In other words, under these circumstances only a dictatorial rule will be g-strategy-proof.

We now review several voting scenarios (procedures, defined over ballots, together with assumptions on preferences) and check whether generalised strategy-proofness is satisfied or violated. Throughout, we assume that lifted preferences are defined in terms of the Gärdenfors axioms.

Recall that the Gibbard-Satterthwaite Theorem does not apply to the case of two candidates. It can also be circumvented for larger numbers of candidates when voters are assumed to have dichotomous preferences. Our next result shows that in fact any voting procedure that can be modelled in terms of the longest-path procedure of Definition 2 and a neutral ballot language (which includes plurality, Borda and approval voting) is strategy-proof when each voter is assumed to have preferences allowing them to only distinguish between "good" and "bad" candidates.

Theorem 6 *For 2-level preferences, longest-path voting is g-strategy-proof wrt. minimal sincerity for any neutral ballot language.*

Proof. Given the longest-path procedure f, for a particular choice of ballot language, and a ballot vector b, we need to show how to construct a sincere ballot b_i' for voter i (with true 2-level preferences p_i) such that $f(b_{-i}, b_i')$ is no worse for i than $f(b)$ according to the Gärdenfors lifting. If b_i is itself sincere, then we are done. Now suppose b_i is not minimally sincere. Then there exist candidates c and c' such that $(c, c') \in R_{b_i}^{\prec} \cap R_{p_i}^{\succ}$. Define a new ballot b_i^* as the result of exchanging the position of c and c' in ballot b_i (this is possible, because the ballot language is assumed to be neutral).

The points awarded to c and c' by voter i will be swapped, and the new points given to c are > 0. Hence, c may drop out of the set of winners (if it has been in in the first place), and c' may enter the set of winners (if it has not been in already) or it may even become the sole winner. Furthermore, as i is assumed to have 2-level preferences, c must be one of its top candidates and c' one of its bottom candidates.

We can iterate the operation that lead from b_i to b_i^* until we arrive at a minimally sincere ballot b_i'. At each step, we may remove one of the bottom candidates from the set of winners and/or we may add one of the top candidates, we may replace the entire set with a singleton consisting of a top candidate, or there may be no change at all. Inspection of the Gärdenfors axioms shows that for any concatenation of such operations, the outcome $f(b_{-i}, b_i')$ will be no worse than $f(b)$. $\square$

Theorem 6 does not hold for quantitative sincerity.[2] However,

for the special cases of plurality, Borda, and approval ballots we do obtain g-strategy-proofness even wrt. quantitative sincerity: for plurality and Borda ballots, the notions of minimal and quantitative sincerity coincide; for approval voting, the optimal ballot is the true (2-level) preference structure.

For the next theorem, "approval voting" is understood to refer to the scenario where $\mathcal{P}$ is the set of total orders, $\mathcal{B}$ is the set of approval ballots, and longest-path voting is used.

Theorem 7 *Approval voting is g-strategy-proof wrt. qualitative sincerity (and thus also wrt. minimal sincerity).*

Proof. A qualitatively sincere ballot b_i' that is not worse than the current ballot b_i can be obtained by considering the voter's preferred candidate among the winners, and by approving this candidate and all those above him in the voter's preference ordering. (If there is just one winner and that winner is the least-preferred candidate for the voter, then we can choose any qualitatively sincere ballot.) $\square$

Theorem 7 cannot be extended to quantitative sincerity. Suppose our voter's preferences are $A \succ B \succ C \succ D$. Then the only quantitatively sincere ballot would be to approve of A and B. But if A and B receive the same number of points from the other voters (and C and D receive no points), then only approving of A would be a better strategy for our voter.

We should stress that for many other voting scenarios, g-strategy-proofness cannot be established. For instance, if the Borda rule is used and voters have preferences with more than two levels, g-strategy-proofness is easily violated. Given the overwhelming prevalence of impossibility results in (classical) voting theory, this is not surprising.

6 Computational Analysis

In the absence of strategy-proofness, protection against manipulation can sometimes be obtained by computational intractability of determining a manipulating ballot [Bartholdi *et al.*, 1989]. Computational complexity also plays a role in our setting. Under a g-strategy-proof voting procedure, there always exists a sincere ballot that is optimal for each voter, but this does not necessarily mean that the voter will actually vote by means of a sincere ballot, because it may be computationally hard to find. It is thus preferable to use voting procedures that are g-strategy-proof and in which it is easy to find an optimal sincere ballot.

Suppose a given voting procedure has been found to be g-strategy-proof for a particular choice of parameters. We can define three *degrees of g-strategy-proofness* in view of how hard it is for a voter to identify a ballot that is both sincere and in their own best interest (knowing that the existence of such a ballot is guaranteed):

(1) *Blind g-strategy-proofness:* Our voter can play optimally and sincerely without requiring any information about the ballots submitted by the other voters (i.e., he can compute his optimal sincere ballot in constant time.)

---

[2]Here is a counterexample: Suppose there are four candidates and the ballot language only allows ballots of the form $\_ \sim \_ \succ \_ \sim \_$ and $\_ \succ \_ \sim \_ \succ \_$. If my true preferences are $A \sim B \succ C \sim D$, then submitting that true preference would be the only quantitatively sincere ballot, while submitting $A \succ B \sim C \succ D$ would yield a better outcome in case D was leading by 2 points over A.

(2) *Tractable g-strategy-proofness:* Our voter needs to know (part of) the ballots of the other voters, but can compute an optimal sincere ballot in polynomial time.

(3) *Intractable g-strategy-proofness:* Our voter needs to know the other ballots and the problem of finding an optimal sincere ballot is computationally intractable.

As a general observation, we note that whenever g-strategy-proofness is known for some notion of sincerity, the number of sincere ballots is polynomial, and winners can be computed in polynomial time, then we get at least tractable (and possibly blind) g-strategy-proofness. A brute-force algorithm could then simply compute the outcome for each sincere ballot and choose the best.

For elections with only two candidates, it is easy to see that the plurality rule is blindly strategy-proof: you should always vote for your preferred candidate, whatever the other ballots are. The same is true for Borda and approval voting. For 2-level preferences over arbitrary numbers of candidates, we can still guarantee tractable strategy-proofness:

Theorem 8 *For 2-level preferences, plurality, Borda and approval voting are all tractably g-strategy-proof wrt. minimal, qualitative, and quantitative sincerity.*

Proof. As shown in the text after Theorem 6, all three rules are g-strategy-proof wrt. all three forms of sincerity. For plurality and approval voting the claim then follows from the fact that the number of sincere ballots is polynomial (for any of the three notions of sincerity). For Borda, the number of sincere ballots is exponential. However, we just need to compute the Borda score of all candidates by the other voters, and then build a Borda ballot (that is, a total order) where all the top level candidates are listed in decreasing order of Borda score, and then the bottom level candidates are listed in increasing order of their score. This ballot is sincere and optimal, and can be found in polynomial time. □

Finally, we refine Theorem 7:

Theorem 9 *Approval voting is tractably g-strategy-proof wrt. minimal and qualitative sincerity.*

Proof. The number of sincere ballots is linear in the number of candidates, so we can simply try them all. □

7   Conclusions and Future Work

We have generalised the traditional setting of voting theory to allow for incomparability and indifference in the preference ordering, as well as for using different languages to model actual preferences and ballots cast in an election. As a generalisation of the concept of truthfulness, we have proposed and analysed three alternative definitions of sincerity. We have then defined a generalised notion of strategy-proofness, and we have proved that some voting procedures are strategy-proof in this more general sense: there is a sincere way of voting which is no worse, for the voter, than any other way of voting. Specifically, we have seen that the classical system of approval voting can be modelled appropriately in our framework and that it is strategy-proof in the generalised sense.

Moreover, we have seen several cases where it is computationally easy for voters to both act sincerely and in their own best interest.

It is likely that our analysis can also be extended to other voting procedures that cannot be modelled as instances of the longest-path voting procedure. An interesting case is STV (*single transferable vote*), for which manipulation is possible but known to be computationally hard in the classical setting [Bartholdi and Orlin, 1991]. It would be interesting to see whether there are scenarios in which a variant of STV is generalised strategy-proof and, if so, to study the complexity of computing an optimal sincere STV ballot.

While strategy-proofness in the classical setting is not possible for non-dictatorial voting procedures, there are some positive results that rely on the restriction of the class of reported preferences. For example, voting with *single-peaked* preferences is strategy-proof in the classical sense [Black, 1958]. Integrating this kind of restriction (defined over preference profiles rather than individual preferences) with the simple structural restrictions studied here constitutes another interesting direction for future work.

References

[Barberà *et al.*, 2004] S. Barberà, W. Bossert, and P. Pattanaik. Ranking sets of objects. In *Handbook of Utility Theory*. Kluwer, 2004.

[Bartholdi and Orlin, 1991] J. J. Bartholdi and J. B. Orlin. Single transferable vote resists strategic voting. *Soc. Choice Welf.*, 8(4):341–354, 1991.

[Bartholdi *et al.*, 1989] J. J. Bartholdi, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Soc. Choice Welf.*, 6(3):227–241, 1989.

[Black, 1958] D. Black. *The Theory of Committees and Elections*. Cambridge University Press, 1958.

[Brams and Fishburn, 2002] S. J. Brams and P. C. Fishburn. Voting procedures. In *Handbook of Social Choice and Welfare*. Elsevier, 2002.

[Endriss, 2007] U. Endriss. Vote manipulation in the presence of multiple sincere ballots. In *Proc. TARK*, 2007.

[Gibbard, 1973] A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.

[Pini *et al.*, 2007] M. S. Pini, F. Rossi, K. B. Venable, and T. Walsh. Incompleteness and incomparability in preference aggregation. In *Proc. IJCAI*, 2007.

[Roberts, 1979] F. S. Roberts. *Measurement Theory*. Addison-Wesley, 1979.

[Satterthwaite, 1975] M. A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *J. Econ. Theory*, 10:187–217, 1975.

[Walsh, 2007] T. Walsh. Representing and reasoning with preferences. *AI Magazine*, 28(4):49–70, 2007.

[Wooldridge, 2002] M. Wooldridge. *An Introduction to Multiagent Systems*. John Wiley & Sons, 2002.