Alice and Bob Will Fight: The Problem of Electing a Committee in the Presence of Candidate Interdependence[1]

Joel Uckelman[2]

Abstract. The problem of electing a committee which satisfies voters is one for which good solutions are scarce. Extending single-winner voting rules to the multi-winner case works well only when voters have no preferential dependencies among candidates for the committee. (Our motivating example is a voter who believes that Alice and Bob are the best candidates, but also that the worst possible committee is one with *both* Alice and Bob.) In order to tackle the interdependence problem, we propose a voting rule called the Goalbase Summation Rule (GSR), which uses goalbases (sets of weighted propositional formulas) as ballots. Using goalbases as ballots lets voters express complex preferences in a compact fashion, while the computational complexity of finding winning committees remains reasonable when the number of seats is fixed. Additionally, the GSR is able to simulate and extend many existing voting rules.

1 Introduction

The problem of electing a committee which satisfies voters is one for which good solutions are scarce. There have been numerous attempts to devise methods for committee election, some which have their origins in single-winner voting methods [4, 5, 6], others of which are intended to produce outcomes which are proportional in some way [7, 18]. Single-winner voting systems frequently fail to respect voter preferences when extended to a multi-winner setting, due mainly to the fact that they deny voters the ability to express interdependence among candidates. Moreover, the way in which such systems measure the "representativeness" of committees may not be at all similar to the way in which voters measure it.

Overcoming preferential dependence is an important issue in decision making, whether the voters are people or software agents. In this latter guise, the problem arises in Multiagent Systems and Computational Social Choice—where there is an extensive literature on preference representation and associated problems (e.g., [8, 9, 10, 13, 14, 15, 16, 17, 20, 21])—and hence solutions for the problem are applicable to AI. To tackle the interdependence problem, we propose a voting method which uses goalbases as ballots, in the spirit of combinatorial vote [15].

In Section 2, we present some known methods for committee selection and find fault with them due to their lack of expressive power. In Section 3 we introduce the Goalbase Summation Rule, which uses sets of weighted propositional formulas as ballots, is able to handle dependencies between candidates in a compact fashion, and can simulate and extend many other voting rules. In Section 4 we consider the computational complexity of finding winning committees using this method. Finally, Section 5 discusses issues arising from the use of the GSR and touches on several avenues for further investigation.

2 Existing Methods for Committee Election

Voters cast ballots for candidates; voting theory concerns itself with the properties of *voting rules*, which map collections of ballots to winners (a subset of the candidates). In what follows, we make frequent reference to four voting rules: *Plurality* is the familiar method where each voter casts a single vote. *Approval* permits a voter to cast at most one vote per candidate. *Borda* has voters ranking the candidates, with each candidate receiving a number of points equal to the number of other candidates who rank below him. *Cumulative* vote gives each voter a supply of points which he may distribute over the candidates as he wishes. In all four cases, the winner is the candidate who receives the most points. There are many other voting rules, but these four suffice for our present purposes.

Various methods for committee elections have been proposed. (For a general discussion of the difficulties of committee elections, see [8] and [16].) The naïve—and perhaps for that reason, most widely used—approach is to extend single-winner plurality to choose more winners.

Definition 1 (m-vote, ℓ-support, top-k)**.** *Call a voting rule* m-vote *if each voter may cast at most m votes;* ℓ-support *if each voter may cast at most ℓ votes per candidate; and* top-k *if the k candidates receiving the most votes are the winners.*

Each tuple $\langle k, \ell, m\rangle$ where $k, \ell, m \in N$ and $\ell \leq m$ defines a voting rule. Standard, single-winner plurality is the 1-vote 1-support top-1 rule, while approval voting is the n-vote 1-support top-1 rule. Cumulative vote is m-vote m-support top-1, while Borda is $\frac{n(n-1)}{2}$-vote $n-1$-support top-1 with the additional constraint that support must be given exactly in the amounts $n-1, n-2, \ldots, 1, 0$.

All m-vote ℓ-support top-k rules have the following flaw when $k > 1$: Voters for whom the candidates are not independent may have no principled way to vote. For example, suppose that we are electing a three-seat committee from the five candidates Alice, Bob, Charlie, Dave, and Elaine (a, b, c, d, e, respectively). Suppose further that a voter Valerie believes that

1. Alice and Bob are the best candidates, so any committee with one of them is better than any committee with neither, and
2. Alice and Bob will fight if they are on the committee together, so any committee with both is worse than any committee with neither.

These two constraints generate the preference order

$$acd, ace, ade, bcd, bce, bde > cde > abc, abd, abe,$$

---

[1] A 2-page version of this paper will appear in the proceedings of ECAI-2010.

[2] Institute for Logic, Language and Computation, University of Amsterdam. Email: `j.d.uckelman@uva.nl`

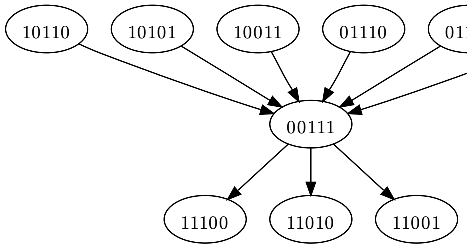

Figure 1. A realistic order on ballots.

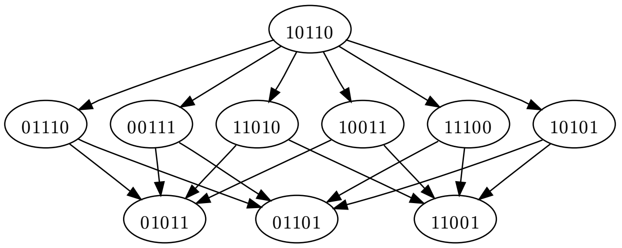

Figure 2. Order on ballots induced by Hamming distance from 10110.

as seen in Figure 1, where the committees are represented as binary vectors and nodes at the same level are equally preferred.

How should Valerie vote if faced with an m-vote ℓ-support top-3 rule? If Valerie votes for both Alice and Bob in hopes that only one will win, she risks electing one of her least-favored committees. If Valerie votes for neither Alice nor Bob in hopes that other voters will prefer one over the other, she risks electing her second choice committee cde. If Valerie votes for Alice but not Bob, or Bob but not Alice, she has no principled way to choose between these options, as she prefers Alice-committees and Bob-committees equally. In fact, Valerie's problem arises for nearly all such rules:

Proposition 2. *Let $a, b, x, y \in C$ (the set of candidates) be distinct, $\vec{c} \subseteq C \setminus \{a, b, x, y\}$, and $|C| \geq 4$. Then no m-vote ℓ-support top-k rule has a ballot compatible with (i.e., induces an order which extends) the partial order $ab\vec{c} < xy\vec{c} < ax\vec{c}, by\vec{c}$, for $k \geq 2$.*

Proof. Fix a ballot $c_1, \ldots, c_n$, where c_i denotes the number of votes given that candidate. Suppose that the ballot induces the order $ab\vec{c} < xy\vec{c} < ax\vec{c}, by\vec{c}$. Wlog, let $\vec{c} = \emptyset$. So $a + b < x + y < a + x,\ b + y$. But then $a + b < a + x \implies b < x$, and $x + y < b + y \implies x < b$, contradiction. □

As summing votes seems to drive this impossibility result, we could try a distance-based vote aggregator such as minisum [4]. Consider ballots and committees as binary vectors; then the winning committee *mini*mizes the *sum* of the Hamming distances to the ballots cast:

$$c \text{ is a winner} \iff \forall c' \in C^k \colon \sum_{b \in B} \mathrm{H}(c, b) \leq \sum_{b \in B} \mathrm{H}(c', b),$$

where C^k is the set of k-seat committees and B is the multiset of ballots cast, expressed as binary vectors. (The Hamming distance H between a ballot and a committee is the least number of bits which must be flipped to transform one vector into the other. For example, $\mathrm{H}(01\underline{010}, 01\underline{101}) = 3$.) Intuitively, any winning committee is one which is as similar in membership as it can be to as many ballots as it can be. This intuition is *wrong*: Any m-vote 1-support k-minisum winner will consist of the top k candidates [6, Appendix, Prop. 4].

Consider instead minimax [6], where

$$c \text{ is a winner} \iff \forall c' \in C^k \colon \max_{b \in B} \mathrm{H}(c, b) \leq \max_{b \in B} \mathrm{H}(c', b),$$

with the variables as before. The m-vote 1-support k-minimax rule is not the same as any summation rule for nontrivial numbers of voters and candidates, but has the quirk that the number of each type of ballot cast is irrelevant and the farthest outlier has the most influence on the result. If one incorrigible voter casts the ballot 00101 while all others choose 11101, then it makes no difference if there are two, ten, or a million ballots cast in total—the winning three-member committees are 10101 and 01101. Usually this is not a desirable feature (though it arguably is appropriate in some circumstances, e.g., multilateral treaty negotiation [4]).

While the *intent* of minimax is to minimize the dissatisfaction of the farthest outlier, the *rule* is to minimize the dissimilarity between the farthest outlier's ballot and the winning committee. But why should we suppose that the farthest outlier (or any voter, for that matter) actually cares about his ballot's similarity to the winning committee? It is quite reasonable to think that some voters' committee preferences will not track the Hamming distance at all. For example, if c is voter v's preferred committee, then any substitution of n members in c is strictly better according to Hamming distance than every substitution of m members, for $n < m$, and all committees c' which are Hamming-equidistant from c are equally preferred by v—both of which are dubious when applied to voters electing real committees.

Returning to our previous example, observe that Valerie's preference ordering is sensitive to small changes in committee composition. Each of the best committees is only one substitution away from some worst committee. Put another way, the Hamming distance between some pairs of best and worst committees is 2, which is always the minimum Hamming distance between two committees of the same size. From Valerie's point of view, $acd \sim bde > abc$; but $\mathrm{H}(10110, 01011) = 4$ while $\mathrm{H}(10110, 11100) = 2$, so the ordering induced by the Hamming distance from acd is $acd > abc > bde$, thus putting an optimal committee last and one of her least favored committees in second place (*cf.* Figures 1 and 2). If we use a minimax procedure and have many voters with preferences like Valerie's, we risk an outcome that is similar in composition to voters' first choices, yet is widely disliked. Furthermore, m-vote 1-support k-minimax rules suffer from the same limited expressivity as top-k rules were shown to in Proposition 2.

The problem we have identified is that the question of committee membership for one candidate is not necessarily independent of the question of committee membership for some other candidate.[3] To borrow from the language of utility functions, some voters have nonmodular preferences. It could be argued that the problem is caused not by the voting methods we examined, but rather because of the way they are applied: The candidates are individuals rather than committees. If committees were raised to the status of first-class citizens—that is, if voters were to vote for whole committees rather than for individuals—perhaps we would not have this problem. However, this approach is unhelpful implemented one way, and scales poorly implemented another. For example, the 2003 Game Theory Society election filled

---

[3] There is a similar problem with the election of representative assemblies—not only might a voter have complex preferences over the composition of the assembly, but preferences over candidates for his district might depend on candidates for other districts where he isn't even able to cast a vote [2].

12 seats from a slate of 24 candidates, giving 2704156 possible committees for voters to consider [6]. This is still manageable if the voter is queried for his most preferred committee only—but then we are left with no information about candidate interdependence, and so we are no better off than before. It is also likely that no two voters will share the same committee as their first choice, so the result will be a many-way tie. If we ask voters for rankings, we begin to sink into the combinatorial morass: Voters would balk at ranking their top 0.001% of the possible committees let alone all 2.7 million of them, and even if the voters were able to rank all of the possible committees the vote tabulator would be overwhelmed by the data.

The unsuitability of single-candidate voting systems for electing committees may also be seen by observing how many voter profiles single-candidate voting methods can accommodate. The most expressive voting method considered above is approval voting, where every subset of candidates is a valid ballot. (All other methods mentioned restrict the set of valid ballots to a proper subset of the powerset of candidates.) In comparison, there are

$$\sum_{i=1}^{\binom{n}{k}} \sum_{j=1}^{i} (-1)^{i-j} \binom{i}{j} j^{\binom{n}{k}}$$

distinct voter profiles over k-seat committees chosen from n candidates, which, for any useful value of n, dwarfs the 2^n distinct approval ballots.[4] Taking the Game Theory Society election as our example, the largest term in the sum is $\binom{24}{12}^{\binom{24}{12}} = 2704156^{2704156}$, which is rather large.[5] Clearly, we need a different approach.

3 Extending Voting Rules By Goalbase Ballots

Rather than electing a committee by voting for individual candidates, we might consider voting for *properties* of the winning committee. What Valerie wants is to indicate her preference for "Alice-XOR-Bob-ness" in the winning committee. Now we present a balloting framework in which this is expressible.

First, we require some definitions, following [21]:

Definition 3 (Goalbases and Utility Functions). *A* goalbase *G is a set of* weighted propositional formulas *(φ, w), where φ is formed from Boolean connectives and propositional symbols from the set $\mathcal{PS}$, and $w \in \mathbb{R}$. A utility function $u\colon 2^{\mathcal{PS}} \to \mathbb{R}$ maps* models *$M \subseteq \mathcal{PS}$ to their values. Each goalbase G generates a utility function u_G such that $u_G(M) = \sum_{(\varphi,w)\in G, M \models \varphi} w$. $\mathrm{Var}(G)$ is the set of propositional symbols occurring in G.*

Imaginge a voting rule where voters submit goalbases with binary weights as ballots. Valerie might then wish to cast $G = \{(\neg a \vee \neg b, 1), (a \dot{\vee} b, 1)\}$ as her ballot. The utility function u_G which this ballot generates,

$$u_G(M) = \begin{cases} 2 & \text{if } a \in M, b \notin M \text{ or vice versa,} \\ 1 & \text{if } a, b \notin M, \\ 0 & \text{if } a, b \in M, \end{cases}$$

induces an ordering on the committees which coincides exactly with Valerie's preference ordering. The models acd and bde score 2, ahead of cde which scores 1, ahead of abc which scores 0.

Here we formulate a general voting rule using goalbases as ballots, in which winners correspond to the optimal states over the sum of all voters ballots:

Definition 4 (Goalbase Summation). *If G, G' are goalbases, then*

$$G \oplus G' = \left\{ \Bigg(\varphi, \sum_{(\varphi,a)\in G} a + \sum_{(\varphi,b)\in G'} b\Bigg) \;\middle|\; \varphi \in \mathrm{For}(G \cup G')\right\}$$

is their formula-wise sum, where $\mathrm{For}(G) = \{\psi \mid (\psi, w) \in G\}$.

Definition 5 (Optimal k-Sized Models). *The set $\mathrm{opt}_k(G)$ is the set of optimal k-sized models according to the goalbase G:*

$$\mathrm{opt}_k(G) = \operatorname*{argmax}_{\substack{M \subseteq \mathcal{PS} \\ |M| = k}} u_G(M).$$

Definition 6 (k-Seat Goalbase Summation Rule). *The* k-Seat Goalbase Summation Rule (GSR) *is the voting rule where voters $v \in \mathcal{V}$ submit goalbases as ballots and the winning committees are the members of* $\mathrm{opt}_k(\bigoplus_{v\in\mathcal{V}} G_v)$.

Our first observation about the GSR is that when $k = 1$, we have a voting rule which generalizes a wide variety of standard single-winner rules in the sense that it accepts all of their ballots (after a bijective mapping to goalbases) as valid ballots itself. For example, plurality ballots correspond to goalbases $\{(c, 1)\}$ where c is the voter's top choice; approval ballots correspond to goalbases $\{(c, 1) \mid c \in \mathcal{A}\}$, where $\mathcal{A}$ is the set of the voter's "approved" candidates. In fact, there is an obvious mapping for all m-vote ℓ-support top-k voting rules:

Proposition 7. *Let $\mathcal{B}$ be a set of m-vote ℓ-support ballots. For each ballot $b = v_{c_1}, \ldots, v_{c_n} \in \mathcal{B}$, represented as a vector of votes for each candidate, define the goalbase $G_b = \{(c_i, v_{c_i}) \mid 1 \leq i \leq n\}$, collecting all $G_b \in \mathcal{G}$. Then k-GSR with ballots $\mathcal{G}$ will choose the same winners as the m-vote ℓ-support top-k rule with ballots $\mathcal{B}$.*

On its face, this is just a baroque way of rewriting the ballots of existing voting rules, so is not interesting by itself; rather, the interest is in what it points to: Voting rules we can *simulate* using GSR we can also *extend*. The problem we identified in Section 2 was that multi-winner extensions of single-winner voting rules have too little expressive power when voters have non-independent preferences. The ballots used by each of these rules are a proper subset of all possible goalbase ballots; therefore, relaxing the restrictions on allowable ballots is a direct way of increasing the expressivity of such rules.

This relaxation of ballot restrictions may be done in both the single- and multi-winner cases. We briefly consider the properties of 1-GSR:

Recall that cumulative vote is the single-winner voting rule where each voter is permitted to distribute some number of votes m over the candidates however he sees fit. (That is, cumulative vote is the m-vote m-support top-1 rule.)

Fact 8. *If no restrictions are placed on ballot goalbases, then* 1*-GSR is cumulative voting without vote limits.*

Cumulative voting without vote limits is not a practical voting method, since for voters it is equivalent to playing the game of which voter can write down the largest number for his favored candidate. (Unrestricted 1-GSR is the finite-vote, finite-support top-1 rule.) However, if we restrict the weights appearing in the goalbase ballots to some range of $\mathbb{N}$, then we have the following correspondence:

---

[4] A set of size n may be partitioned into k nonempty subsets $\left\{{n \atop k}\right\}$ ways (where $\left\{{n \atop k}\right\} = \frac{1}{k!}\sum_{j=1}^{k}(-1)^{k-j}\binom{k}{j}j^n$ denotes a Stirling number of the second kind [11, Section 6.1]), and in each case these k subsets may themselves be strictly ordered in $k!$ ways. Thus the number of (not necessarily strict) linear orders on n items is $\sum_{k=1}^{n} k!\left\{{n \atop k}\right\} = \sum_{k=1}^{n}\sum_{j=1}^{k}(-1)^{k-j}\binom{k}{j}j^n$.

[5] For comparison, an estimated 106 billion people had ever lived as of the year 2002 [12]. The profile space is more than adequate to permit every person who has ever lived a unique opinion on the 2003 Game Theory Society committee election.

Fact 9. *If each goalbase G_i uses weights from $0, \dots, n \in \mathbb{N}$ only and $\sum_{(\varphi,w)\in G_i} w \leq m$, then 1-GSR corresponds to m-vote cumulative voting.*

The following is an example of how relaxing restrictions on goalbases can yield new voting rules:

Example 10. Call Property Approval Voting (PAV) the voting rule in which *properties of the outcome* (rather than individual candidates) are the objects of approval or disapproval. Any goalbase G having 1 as the sole permissible weight for any formula constitutes an admissible PAV ballot. However, some formulas will be useless: Any formula which implies a conjunction of positive literals longer than the intended number of winners, and any formula which implies a conjunction of negative literals longer than the intended number of losers, will effectively be equivalent to $\bot$. A significant difference between AV and PAV is the range of preorders of which they permit representation. Every AV ballot induces a dichotomous order, while PAV supports much more. In the case where there are three candidates a, b, c, the PAV ballot $\{(a, 1), (a \vee b, 1)\}$ induces the (non-dichotomous) order $a > b > c$, since the state $\{a\}$ receives two points, $\{b\}$ one point, and $\{c\}$ zero points. (Only singleton states are relevant here, since we are considering the single-winner case.)

In fact, there is a general way of representing any strict linear order $a_1 > a_2 > \cdots > a_n$ with a PAV ballot:

$$\begin{gathered}(a_1 \vee \cdots \vee a_{n-2} \vee a_{n-1}, 1)\\ (a_1 \vee \cdots \vee a_{n-2}, 1)\\ \vdots\\ (a_1 \vee a_2, 1)\\ (a_1, 1)\end{gathered}$$

The clause which ends with a_i is the one which causes a_i to be ordered strictly above a_{i+1}, so by omitting that clause we can get a ballot where $a_i \sim a_{i+1}$. This is sufficient to induce any weak linear order over the candidates. Thus, in the single-winner case, PAV is a variant of the Borda rule.

4 The Complexity of Deciding Winning Slates

If we are to use goalbases as ballots, then in the multi-winner case finding winning committees is an optimization problem. The problem of deciding whether some utility threshold may be reached given a goalbase representation of a utility function arises in other applications of goalbases (e.g., when using goalbases as combinatorial auction bidding languages as in [20, Chapter 6]). Following [21], we give a formal definition of this decision problem, known as MAX-UTIL:[6]

Definition 11 (MAX-UTIL)**.** *The decision problem* MAX-UTIL *has as input a goalbase G and $K \in \mathbb{Z}$, and outputs 1 if some $M \in 2^{\mathcal{PS}}$ has $u_G(M) \geq K$, and 0 otherwise.*

Determining the winner of an election where goalbases are ballots is related to solving MAX-UTIL for the sum of those goalbases. MAX-UTIL for goalbase languages with straightforward definitions tends to be either trivial or NP-complete (see [21, Section 5] and [20, Section 5.5]).

Because we are concerned here with electing committees of a size fixed prior to the election (as opposed to the open-ended committees discussed by Brams *et al.* [6]), we cannot apply MAX-UTIL directly to the sum of voters' goalbases in order to determine the winners of the election. Doing that might yield a model with the wrong number of winners. We must do something to ensure that only models which fill k seats are potentially optimal. One approach to adapting our winner determination problem to MAX-UTIL is to augment the sum of the voters' goalbases with formulas which increase the utility of k-sized models (or decrease the utility of non-k-sized models).

First, some notation is required. Define the following formulas:

$$\begin{aligned}\varphi_{\geq k} &= \bigvee\Big\{\bigwedge X \;\Big|\; X \subseteq \mathcal{PS} \text{ and } |X| = k\Big\}\\ \varphi_{\leq k} &= \neg\varphi_{\geq k+1}\\ \varphi_{=k} &= \varphi_{\leq k} \wedge \varphi_{\geq k}\end{aligned}$$

and the quantity

$$\delta = \sum\Big\{|w| \;\Big|\; (\varphi, w) \in \bigoplus_i G_i\Big\},$$

where G_i is the goalbase of voter i. The formula $\varphi_{=k}$ is such that a model $M \models \varphi_{=k}$ iff $|M| = k$. The quantity δ is a (not necessarily tight) upper bound on utility change between arbitrary models for $u_{\bigoplus_i G_i}$. Note that $\varphi_{\geq k}$ has $\binom{n}{k}$ disjuncts, and so is potentially a very long formula.[7] However, in the context of committee elections n and k—the numbers of candidates and seats—will tend to be small, and, as will be seen below, $\varphi_{=k}$ appears exactly once in the goalbase which represents the voters' preferences.

Suppose that $\bigoplus_i G_i$ is the sum of voter goalbases in a k-seat committee election. Let

$$G = \Big(\bigoplus_i G_i\Big) \oplus \{(\varphi_{=k}, \delta + 1)\}.$$

Since δ is an upper bound on utility change between models for $u_{\bigoplus_i G_i}$, we can say the following: If M, N are models such that $M \models \varphi_{=k}$ and $N \not\models \varphi_{=k}$, then $u_G(M) > u_G(N)$, as the greatest possible utility loss of moving from N to M in $u_{\bigoplus_i G_i}$ is δ, and making $\varphi_{=k}$ true results in a gain of $\delta + 1$. Thus, since any model of size k is strictly better than every model of any other size, we are guaranteed that all models which yield maximal utility are of size k. Moreover, since $\varphi_{=k}$ is true on *every* model of size k, it does not affect their utility relative to one another, so augmenting $\bigoplus_i G_i$ with $(\varphi_{=k}, \delta + 1)$ preserves the ordering of (relevant) models.

Therefore, we may easily adapt the input to force size-k models to the top of the ordering and use an off-the-shelf algorithm for deciding MAX-UTIL to determine winners—though this may be impractical due to the complexity of MAX-UTIL for unrestricted goalbases. If $\bigoplus_i G_i$ is confined to something less than the full goalbase language, however, we may be able to use that to our advantage. The following proposition shows that when solving MAX-UTIL we may always reduce a goalbase outside a given language $\mathcal{L}$ to a goalbase inside the language by solving a simpler version of MAX-UTIL at most an exponential number of times:

---

[6] In [20, 21] a parametrized version of MAX-UTIL is used. We treat the parameters informally here.

[7] While it is not possible to shorten $\varphi_{\geq k}$ using standard Boolean connectives, we can be more concise if we are willing to augment our language with a cardinality operator. For example, Benhamout *et al.* [1] consider a variant of propositional logic in which there are *pair formulas* $(\rho, \mathcal{L})$, where $\mathcal{L}$ is a multiset of literals and ρ specifies how many elements of the multiset must be true in order for the $(\rho, \mathcal{L})$ to be true. Clearly $(|\mathcal{PS}|/2, \mathcal{PS})$ is equivalent to $\varphi_{\geq|\mathcal{PS}|/2}$, but exponentially shorter. Hoos and Boutilier [13] propose a similar, though less powerful, k-of operator—less powerful due to the fact that their (bidding) language lacks negation, and so any k-of operates on atoms only.

Proposition 12. *If $G \in \mathcal{L}$, $G' \notin \mathcal{L}$, and $\mathcal{L}$ is closed under substitution of logical constants for atoms, then* MAX-UTIL *for $G \oplus G'$ can be solved with $\leq 2^{|\mathrm{Var}(G')|}$ calls to a* MAX-UTIL *oracle for $\mathcal{L}$.*

Proof. There are $2^{|\mathrm{Var}(G')|}$ models on just the variables occurring in formulas in G'. For each model over the variables in G', we substitute $\top$ and $\bot$ into $G \oplus G'$ as per the model and carry out MAX-UTIL on the modified $G \oplus G'$. □

If $\mathrm{Var}(G')$ is small and does not depend on $\mathcal{PS}$, decomposing a goalbase containing alien formulas in this way is potentially feasible. However, the formula $\varphi_{=k}$ contains every atom in $\mathcal{PS}$ at least once and hence the upper bound we get is exponential in $|\mathcal{PS}|$, which is unhelpful. (For a discussion of the substitution closure condition, see [20, Section 5.7.1].)

An alternative approach is to modify the decision problem instead of the goalbase. Perhaps MAX-UTIL is not the right decision problem unless we have the same number of candidates as seats—in which case, why vote? Instead, we define a variant of MAX-UTIL where exactly k atoms must be true in any solution:

Definition 13 (k-MAX-UTIL)**.** *The decision problem k-*MAX-UTIL *has as input a goalbase G and $K \in \mathbb{Z}$, and outputs 1 if some $M \in \mathrm{opt}_k(G)$ has $u_G(M) \geq K$, and 0 otherwise.*

k-MAX-UTIL is the decision-problem version of finding members of $\mathrm{opt}_k(G)$, just as MAX-UTIL is the decision-problem version of finding members of $\mathrm{opt}(G) = \bigcup_{i=0}^{|\mathcal{PS}|} \mathrm{opt}_i(G)$.

Fortunately, having a fixed number of seats to fill dramatically reduces the complexity of finding a voter's preferred ballot:

Proposition 14. *k-*MAX-UTIL *$\in$ P, for fixed $k \in \mathbb{N}$.*

Proof. For any given k and $\mathcal{PS}$, there are $\binom{|\mathcal{PS}|}{k}$ models of size k to check. It is always the case that $\binom{n}{k} \leq \frac{n^k}{k!}$, which grows polynomially in $n = |\mathcal{PS}|$ for any fixed k. □

This makes *whatever* goalbase language we want for representing our voters' ballots computationally tractable (though not necessarily trivial) so long as the number of seats and candidates is not too large. In particular, it is well within the capabilities of contemporary desktop computers to determine the winners in committee elections of a size similar to that conducted by Game Theory Society in 2003 (see Section 2), where there would be only 2.7 million models to check. In practice, an off-the-shelf integer program solver could be used to find winners, since any set of ballots can be converted to an integer program.

5 Discussion & Future Work

Many paths are yet to be explored. In this section we give an overview of those of which we are aware.

In order to use goalbase ballots for multi-winner voting, we must place some restriction on the weights which are available to voters. As noted after Fact 8, cumulative voting without point limits is not a sensible voting method. Having established that restrictions are needed, we are faced with the problem of selecting some—it is not presently obvious which restrictions are most suitable. The restriction which cumulative voting itself suggests is to limit the sum of weights in any goalbase: $\sum_{(\varphi,w)\in G} w \leq K$.[8] This is a limit on the input space. Another approach is to restrict the output space: For example, we might limit the utility of any admissible state: $u_G(M) \leq K$ for all $M \subseteq \mathcal{PS}$ where $|M| = k$.

There are advantages and disadvantages to both methods. If our voter is a person, he will find it easier to cast a valid sum-limited ballot than a valid state-limited one. Input limits are not uncommon: For example, in the U.S., the State of Illinois used single-winner cumulative voting with a 3-point limit (i.e., 3-vote 3-support top-1) within districts for electing members of its House of Representatives from 1870 to 1980 [19, 22]. Corporate boards of directors are usually elected using cumulative voting, where the point limit for each voter is the number of shares he owns. We know of no uses of output limits. Presumably this is because it is hard to see when working in the input space whether output limits are being respected; output limits expect too much of the average voter. However, output limits on elections of the size human voters are likely to face will not be difficult for machines to enforce, so might be practical if voters use a computer-aided voting system. This is a user-interface issue.

Point limits also raise a fairness issue. For simplicity, we use a single-winner example, though the problem it illustrates is general. The sum-limit $\sum_{(\varphi,w)\in G} w \leq K$ will not always produce utility functions having equal sums for singleton models. E.g., consider the goalbase ballots $G_1 = \{(a, 10)\}$ and $G_2 = \{(a \vee b, 10)\}$. The latter has a singleton state sum of 20 ($u_{G_2}(\{a\}) = 10, u_{G_2}(\{b\}) = 10$), while singleton states for the former sum only to 10 ($u_{G_1}(\{a\}) = 10$). The sum-limit gives voters with top-heavy preferences more influence on the outcome than voters with balanced or bottom-heavy preferences. We could try to mitigate this by "normalizing" formulas based on the number of states they affect, e.g., $(a \vee b, 10)$ could be translated to $(a, 5), (b, 5)$, but this would seem to disadvantage voters who have top-heavy preferences. If $(\bigvee \mathcal{PS} \setminus \{a\}, 10)$, after normalization, gives one point to everyone but candidate a, that is not likely to be very effective for voters who dislike a but otherwise do not distinguish among the other candidates. Or we could try other ways of normalizing—Lafage and Lang [14, Section 3.2.3] suggest postprocessing (dis)utilities to equalize entropy across agents—which will potentially have some other differential effect on voters.

The basic question here seems to be how to set the value of preferences which are not over single states against those which are. What is an appropriate measure of voting power here? Input limits seem to favor top-heavy voters, output limits seem to favor bottom-heavy voters. One way of quantifying the effect that a proposed weight limit could have is by considering the *efficacy* of voters with different preferences under that weight limit. (The efficacy of a ballot for a voter is a measure of how often that voter will be pivotal if he casts that ballot.) Ideally, all voters would have equally efficacious ballots to cast. Brams and Fishburn [3, Chapter 5] calculate the efficacy of ballots for approval voting and find that not all ballots are equally effective. If we assume that our voters are truthful, what this means is that approval voting is advantageous for voters with some kinds of preference orders and disadvantageous for others. A similar analysis could be done for cumulative voting with goalbase ballots, with an eye to which weight restrictions treat voters most equitably.

With any voting system, there are questions about whether it encourages or discourages strategic voting. The manipulability of a voting system must always be considered in the context of a notion of sincerity, for we cannot say whether a voter is *mis*representing his preferences if we cannot first say what it would be for a voter to represent his preferences accurately.

Consider, first, voting systems with ordinal ballots. Many standard systems—e.g., plurality, approval, Borda—use ballots which contain

---

[8] If we permit negative weights, then we would need to place an upper bound on the sum of the absolute values of the weights instead of on the sum of the weights. In this way we avoid ballots like $\{(a, -2^{1000}), (b, 2^{1000} + 10)\}$ which the voter could otherwise claim is a 10-point ballot.

purely ordinal information. In the case where there is an allowable ballot which induces the same preorder over outcomes as the voter's true preorder, then any reasonable notion of sincerity should deem that ballot sincere (and any ballot which does not induce that same preorder, insincere). This means, for example, that for voters with strict linear orders, there will always be unique sincere plurality and Borda ballots; and similarly, for voters whose preferences are dichotomous (and not wholly indifferent), there will be a unique sincere approval ballot. However, this will not be the case for voters with other kinds of preorders. There are no approval ballots which express nondichotomous preferences (e.g., $x > y > z$); standardly, Borda does not permit ties, so voters with weak (instead of strict) orders will have no ballots which express their preferences exactly.

What should count as sincere in the space of cardinal ballots is not immediately obvious. A voter's preferences may be inexpressible as a result of restrictions on the ballot language, and this can result in the existence of multiple sincere ballots which the voter could cast. Endriss [9] explores the existence of multiple sincere ballots for approval voting and shows that the Gibbard-Satterthwaite Theorem is avoidable in that context; Endriss *et al.* [10] present several measures of sincerity for languages where ballots are preorders, and examine the consequences for strategyproofness under these. This line of research could be continued for goalbase ballots, first by developing reasonable notions of sincerity, and secondly by determining which language restrictions induce sincerity in rational voters. Meir *et al.* [17] avoid the problem of sincerity in multi-winner voting altogether by defining manipulation as an optimization problem asking whether, given the ballots of some other voters, there is a ballot which the manipulating voter may cast which yields him at least t utility. The question of whether a better ballot exists is more general than, and serves as a proxy for, the question of whether a better insincere ballot exists—though this still leaves open the possibility that some ballot which is optimal is nonetheless also sincere, and so does not exactly capture classical manipulability.

Finally, we might consider questions about the difficulty of finding a sincere ballot given a voter's preferences. It would not be surprising to learn that for some languages, it is always in the voter's best interests to cast a sincere ballot, but nonetheless quite difficult for him to determine which ballots are sincere for him. Strategyproofness is not worth much in this case. A method for constructing sincere ballots will be essential for any language intended for human voters.

6 Conclusion

In this paper we considered some methods for electing committees and demonstrated that they lack properties which are desirable when conducting multi-winner elections. In particular, single-winner voting methods lack the expressivity to extend well to the multi-winner case. The observation that it is possible to simulate many single-winner voting methods using the Goalbase Summation Rule, which takes goalbases as ballots, suggests one way of extending the expressivity of existing voting methods for use in a multi-winner setting. Because multi-winner elections tend to have the number of winners fixed beforehand, the complexity of determining winners (using the decision problem MAX-UTIL) is limited, even when the goalbase language is unrestricted. Along these lines, we suggest as an example a multi-winner extension of approval voting, which we call Property Approval Voting. Finally, we discuss the possibilities for future work: the need to find useful limits on weights in goalbase ballots; the fairness of these limits, since they may differentially affect voters with dissimilar preferences; and issues related to sincerity and strategic voting.

ACKNOWLEDGEMENTS

This work is not entirely disjoint with [20, Chapter 7], which was supported by a GLoRiClass fellowship funded by the European Commission (Early Stage Research Training Mono-Host Fellowship MEST-CT-2005-020841).

REFERENCES

[1] B. Benhamou, L. Sais, and P. Siegel, 'Two proof procedures for a cardinality based language in propositional calculus', in *STACS 94, 11th Symposium on Theoretical Aspects of Computer Science*, eds., P. Enjalbert, E. W. Mayr, and K. W. Wagner, pp. 71–82. Springer, (1994).

[2] J.-P. Benoît and L. A. Kornhauser, 'Social choice in a representative democracy', *American Political Science Review*, **88**(1), 185–192, (1994).

[3] S. J. Brams and P. C. Fishburn, *Approval Voting*, Springer, second edn., 2007.

[4] S. J. Brams, D. M. Kilgour, and M. R. Sanver, 'A minimax procedure for negotiating multilateral treaties', in *Reasoned Choices: Essays in Honor of Academy Professor Hannu Nurmi*, ed., M. Wiberg, 108–139, Finnish Political Science Association, Turku, Finland, (2004).

[5] S. J. Brams, D. M. Kilgour, and M. R. Sanver, 'How to elect a representative committee using approval balloting', in *Mathematics and Democracy*, Studies in Choice and Welfare, 83–95, Springer, (2006).

[6] S. J. Brams, D. M. Kilgour, and M. R. Sanver, 'A minimax procedure for electing committees', *Public Choice*, **132**(3–4), 401–420, (2007).

[7] J. R. Chamberlin and P. N. Courant, 'Representative deliberations and representative decisions: Proportional representation and the Borda rule', *The American Political Science Review*, **77**(3), 718–733, (1983).

[8] Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet, 'Preference handling in combinatorial domains: From AI to social choice', *AI Magazine*, **29**(4), 37–46, (2008).

[9] U. Endriss, 'Vote manipulation in the presence of multiple sincere ballots', in *Proc. 11th Conf. on Theoretical Aspects of Rationality and Knowledge (TARK-2007)*, ed., D. Samet, pp. 125–134. Presses Universitaires de Louvain, (2007).

[10] U. Endriss, M. S. Pini, F. Rossi, and K. B. Venable, 'Preference aggregation over restricted ballot languages: Sincerity and strategy-proofness', in *Proc. 21st Intl. Joint Conf. on Artificial Intelligence (IJCAI-2009)*, ed., C. Boutilier, pp. 122–127, (2009).

[11] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, Addison-Wesley, second edn., 1994.

[12] C. Haub, 'How many people have ever lived on earth?', *Population Today*, **30**(8), 3–4, (November/December 2002).

[13] H. H. Hoos and C. Boutilier, 'Solving combinatorial auctions using stochastic local search', in *Proc. 17th Conf. on Artificial Intelligence (AAAI 2000)*, pp. 22–29. AAAI Press, (2000).

[14] C. Lafage and J. Lang, 'Logical representation of preferences for group decision making', in *Proc. 7th Intl. Conf. on Principles of Knowledge Representation and Reasoning (KR-2000)*, eds., A. G. Cohn, F. Giunchiglia, and B. Selman, pp. 457–468. Morgan Kaufmann, (2000).

[15] J. Lang, 'Logical preference representation and combinatorial vote', *Annals of Mathematics and Artificial Intelligence*, **42**(1–3), 37–71, (2004).

[16] J. Lang and L. Xia, 'Sequential composition of voting rules in multi-issue domains', *Mathematical Social Sciences*, **57**(3), 304–324, (2009).

[17] R. Meir, A. D. Procaccia, J. S. Rosenschein, and A. Zohar, 'Complexity of strategic behavior in multi-winner elections', *Journal of Artificial Intelligence Research (JAIR)*, **33**, 149–178, (2008).

[18] B. L. Monroe, 'Fully proportional representation', *The American Political Science Review*, **89**(4), 925–940, (1995).

[19] B. F. Moore, *The History of Cumulative Voting and Minority Representation in Illinois, 1870–1908*, volume III.3 of *The University Studies*, University Press, Urbana-Champaign, Illinois, 1909.

[20] J. Uckelman, *More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains*, Ph.D. dissertation, University of Amsterdam, 2009. ILLC Publication DS-2009-12.

[21] J. Uckelman, Y. Chevaleyre, U. Endriss, and J. Lang, 'Representing utility functions via weighted goals', *Mathematical Logic Quarterly*, **55**(4), 341–361, (2009).

[22] 'Alternative voting systems as remedies for unlawful at-large systems', *The Yale Law Journal*, **92**(1), 144–160, (1982).