Electing a committee with constraints

Egor Ianovski

Abstract

We consider the problem of electing a committee of \(k\) candidates, subject to some constraints as to what this committee is supposed to look like. In our framework, the pool of candidates is divided into tribes, and constraints of the form “at least \(p\) candidates must be elected from tribe \(X\)” and “there must be at least as many members of tribe \(X\) as of \(Y\)” are considered. While in general this problem would require us to rethink how we determine which election outcomes are good, in the case of a committee scoring rule this becomes a constrained optimisation problem – simply find a valid committee with the highest score. In the case of weakly separable rules we show the existence of a polynomial time solution in the case of tree-like constraints, and a fixed-parameter tractable algorithm for the general case, which is otherwise NP-hard.

1 Introduction

Perhaps the least controversial desideratum in social choice theory is non-imposition – the requirement that every candidate can be a winner in at least one profile. Indeed, it is hard to come up with a convincing story why an election designer should allow voters to vote for \(a\), while eliminating even the theoretical possibility of \(a\) winning, unless he is actively trying to provoke a revolution.

The situation changes when we consider multiwinner elections. When electing assemblies, parliaments, and committees (or, indeed, “electing” a movie library or a package of advertisements) we often encounter constitutional or conventional restrictions on which sets are acceptable and which are not. This could be due to equity concerns, such as the twenty-four countries around the world that reserve seats for women; protection of minority rights, such as the religious seats in Iran or the ethnic seats in Croatia; social stability, such as the Columbian peace agreement which reserved seats for former FARC combatants, or the Roman requirement that one consul be a pleb; credibility, such as the bipartisan committees in the United States, or the Cypriot Supreme Court which requires a Greek, Turkish, and a neutral judge; protection of culture, such as the French law requiring that forty percent of songs sung on radio are in French; and many others.
This creates difficulties for social choice theory because such constraints are exogenous to the standard framework. A multiwinner election typically has access to a list of candidates, and the voters’ preferences over them. The candidates are just a list of names, and any sensible function will treat them symmetrically. The function does not have access to the fact that failing to elect a will cause the army to secede and let the barbarians through the gates. Even if we have access to such constraints, however, the remains the perennial problem of social choice – out of all the committees that satisfy the constraints, clearly we want the best one. But what does that mean?

In this paper we will make a start on this problem by suggesting a framework for specifying constraints on committee composition, and considering how such a committee may be elected under a given committee scoring rule. The advantage of using a scoring rule in this instance is that it offers a clear answer to how to determine which election outcomes are desirable – if we believe that the score produced by the voting rule is indeed a reasonable measure of social welfare, then the problem is simply to find a committee satisfying the constraints which maximises this score.

1.1 Related work

Committee scoring rules were first introduced by [4], in which the authors identify the classes of weakly separable and representation-focused rules, and study the properties committee selection rules might be expected to satisfy with respect to three possible applications. Weakly separable rules are found to be tractable for reasonable underlying single-winner functions, while representation-focused rules in general are NP-hard, following from the results of [14, 2, 13].

A third class, the top-\(k\) counting rules, was introduced by [7] in the context of finding a multiwinner analogue of the fixed-majority criterion. The superclass of ordered weighted average rules was introduced by [15], and the relationship between these classes and their axiomatic properties was studied by [6].

The notion that the outcome of a multiwinner election may be restricted to a set of admissible committees is not in itself new. It is part of the framework of [10, 11] in the context of approval voting, though in their work the focus is on restricting a general multiwinner election to one that elects a set of fixed size – in our parlance, a committee selection rule – and they do not consider the algorithmic aspects of electing a committee with more general restrictions; and it is interesting that the Duggan-Schwartz theorem ([3, 16]), in extending the Gibbard-Satterthwaite theorem to multivalued elections, does not include surjectivity as one of its hypotheses. However, to my knowledge there has been no work done on actually considering the problem of how to actually elect a committee from a restricted admissible set.

2 Preliminaries

2.1 Committee scoring rules

Let \([m]\) be the set of integers \(\{1, \ldots, m\}\), and \([m]_k\) be the set of all length-\(k\) increasing sequences of numbers from \([m]\). Given two sequences, \(I = (i_1, \ldots, i_k)\) and \(J = (j_1, \ldots, j_k)\), the score of sequence \(I\) under rule \(R\) is denoted \(R(I)\). The goal is to find a committee \(C\) that maximises \(R(C)\) subject to \(C \in [m]_k\).
(j_1, ..., j_k), we write I \succeq J if for each t \in [k], it holds that i_t \geq j_t.

An election, E, is a triple \((C, V, k)\) consisting of a set of candidates, C, a set of voters, V, and a committee size k. Every voter is identified with a linear order over C, which we call a preference order. We use pos_{v_i}(c), c \in C, to denote the position of c in v_i’s preference order. A committee selection rule is a function which takes an election to a subset of candidates of size k, which we call a k-committee.

For a k-committee X, we use pos_{v_i}(X) to denote the sequence that we obtain by sorting the set \(\{\text{pos}_{v_i}(c) \mid c \in X\}\) in increasing order. Naturally, \(\text{pos}_{v_i}(c) \in [m]\) and \(\text{pos}_{v_i}(X) \in [m]_k\).

The class of committee selection rules we are interested in operates by assigning a number of points to each committee for each voter, where the number of points assigned to X for voter i is a function of \(\text{pos}_{v_i}(X)\).

**Definition 2.1.** A committee scoring function for m candidates and committee size k is a function \(f_{m,k} : [m]_k \rightarrow \mathbb{R}^+\) such that, for any sequences I, J \in [m]_k, if I \succeq J then \(f_{m,k}(I) \geq f_{m,k}(J)\).

Let \(f = (F_{m,k})_{k \leq m}\) be a family of committee scoring functions. The induced committee scoring rule is a function \(R_f\) that given an election E and an integer k outputs all k-committees that maximise \(\text{score}_E(X) = \sum_{v_i \in V} f_{m,k}(\text{pos}_{v_i}(X))\).

[6] identify a hierarchy of such rules. They get hard very quickly: of the three classes at the bottom – weakly separable, top-k counting, representation-focused – only weakly separable rules are polynomial-time computable in the general case,\(^1\) and the top-k counting and representation-focused rules known to be easy are thus because of their similarity to weakly separable rules. Since our focus in this paper is computational, we will only concern ourselves with separable rules.

**Definition 2.2.** We say that a family of committee scoring functions \(f = (f_{m,k})_{k \leq m}\) is weakly separable if there exists a family of single-winner scoring functions \((\gamma_{m,k})_{k \leq m}\) with \(\gamma_{m,k} : [m] \rightarrow \mathbb{R}^+\) such that for every \(m \in \mathbb{N}\) and every \((i_1, ..., i_k) \in [m]_k\) we have:

\[
f_{m,k}(i_1, ..., i_k) = \sum_{t=1}^{k} \gamma_{m,k}(i_t).
\]

A committee scoring rule \(R_f\) is weakly separable if it is defined through a family of weakly separable scoring functions \(f\).

Note that weakly separable rules, as the name suggests, allow the score of a committee to be separated. By this we mean:

\[
\text{score}_E(X) = \sum_{v_i \in V} f_{m,k}(\text{pos}_{v_i}(X))
= \sum_{v_i \in V} \sum_{j \in X} \gamma_{m,k}(\text{pos}_{v_i}(j))
= \sum_{j \in X} \sum_{v_i \in V} \gamma_{m,k}(\text{pos}_{v_i}(j)).
\]

\(^1\)Subject to assumptions about the underlying scoring functions being polynomial-time computable.
We will thus refer to \( \sum_{v_i \in V} \gamma_{m,k}(pos_{v_i}(j)) \) as the score of \( j \), or \( score_E(j) \).

Natural examples of weakly separable rules are those where the underlying scoring rules are the familiar scoring rules of social choice theory. For example, Single Non-Transferable Vote (SNTV) is the committee scoring rule with the plurality underlying scoring function, \( \gamma_{m,k}(1) = 1, \gamma_{m,k}(i \neq 1) = 0 \); Borda count is the committee scoring rule derived from the Borda function, \( \gamma_{m,k}(i) = m - i \); and Bloc is the rule derived from \( k \)-approval, \( \gamma_{m,k}(i) = 1 \) for \( i \leq k \).

2.2 Range restrictions

In the most general sense, a restriction on the range of a committee selection rule would take the form of some set \( S \subseteq 2^C \) of viable committees, with the requirement that the rule always output a member of \( S \). However from a computational point of view such an approach is neither tenable nor interesting – if \( S \) is large, then listing the admissible sets as input is impractical; if \( S \) is small, then the problem of finding the highest scoring committee can be trivially solved by trying every committee in \( S \). An alternative approach could be to describe \( S \) as a formula in some logical language, \( \varphi_S \). This solves the problem of triviality and input size, but if the language is rich enough to capture propositional logic then the satisfaction problem will already be \( \text{NP} \)-hard, and we will have hit a wall before we even started.

Moreover, it does not seem to me that we need such a level of generality at all. All the constraints used in practice seem to be either reserving seats (i.e. at least a certain number of a group must be elected), partitioning seats (exactly a certain number must be elected), or establishing that one group is not disadvantaged with respect to another. Four constraints seem to capture everything we need:

- \( \text{AtLeast}(p, X) \): at least \( p \) members of the committee are from \( X \).
- \( \text{Exactly}(p, X) \): exactly \( p \) members of the committee are from \( X \).
- \( \text{AsMany}(X, Y) \): at least as many members of the committee are from \( X \) as from \( Y \).
- \( \text{SameNumber}(X, Y) \): the same number of members of the committee are from \( X \) as from \( Y \).

Note that \( \text{SameNumber}(X, Y) \) can be expressed as the pair \( \text{AsMany}(X, Y) \) and \( \text{AsMany}(Y, X) \).

We will keep this constraint for this section because in Proposition 2.1 we will show that it offers us no escape from computational complexity, but to simplify our proofs in the next section we will assume that only \( \text{AsMany} \) constraints are used. Likewise, we introduce a new constraint to allow us to get rid of \( \text{Exactly}(p, X) \):

- \( \text{AtMost}(p, X) \): at most \( p \) members of the committee are from \( X \).

We can now define the algorithmic problems of interest.

**Definition 2.3.** The constrained winner election problem for a committee scoring rule \( R \) is the problem that takes an election \( E \), a set of constraints \( C \), and a set of tribes
\[ T = T_1, \ldots, T_p \subseteq C \] as input. The output is some \( k \)-committee \( X \) that maximises the score out of all the committees that satisfy \( C \).

The constrained winner existence problem for a committee scoring rule \( \mathcal{R} \) is the problem that takes an election \( E \), a set of constraints \( C \), a set of tribes \( T \), and a target score \( S \) as input. The output is YES if there exists a \( k \)-committee \( X \) that satisfies \( C \) and has score at least \( S \), and NO otherwise.

Unfortunately, if \( T \) is arbitrary then simply determining whether there exists a committee satisfying \( C \) is NP-hard, even if we restrict ourselves to any one of the constraints we have outlined above.

Proposition 2.1. It is NP-hard to determine whether there exists a committee \( X \) satisfying a set of constraints \( C \), even if \( C \) consists of only one type of constraint. Thus the constrained winner existence problem is NP-complete for any onto committee selection rule.

Proof. Given a graph \( G = (V, E) \) and an integer \( k \), we can introduce a candidate for every vertex, and encode vertex cover with the constraints \( \{ \text{AtLeast}(1, e_i) \mid e_i \in E \} \), then attempt to elect a committee of size \( k \). With AtMost constraints, the constraints have the same form but the committee is of size \(|V| - k\), representing the vertices not on the cover. If we only have AsMany, we can use \( \{ \text{AsMany}(X, \{c_i\}) \mid c_i \in C \} \) to encode AtLeast(1, \( X \)), and then do the same encoding.

Given a set \( S \) and subsets \( S_i \) of \( S \), we can introduce a candidate for every set and encode exact cover with \( \{ \text{Exactly}(1, \{S_j \mid s_i \in S_j\}) \mid s_i \in S \} \). If we only have SameNumber, we will first introduce constraints \( \text{SameNumber}(\{S_j \mid s_i \in S_j\}, \{S_j \mid s_i \neq s_j \in S_j\}) \) for \( s_i, s_j \in S \). This will ensure that every element is covered by the same number of sets. To ensure that this number is one, we will introduce a new candidate \( c \), add the constraint \( \text{SameNumber}(\{S_j \mid s_i \in S_j\}, \{c\}) \), and elect a committee of size \( k + 1 \).

We will thus need to consider restricted cases for the list of tribes. We will consider the two natural cases of disjoint tribes, and a small number of tribes.

3 Disjoint tribes and knapsack

We note that the \( \text{AsMany}(X, Y) \) constraints impose an ordering on the tribes – if we imagine we are building the committee from the ground up we will have to take a candidate from \( X \) before we take a candidate from \( Y \). If there are several tribes dominating \( Y \) in this order, we will have to take candidates from all of them. As such the problem is reminiscent of partial-order knapsack where, in addition to weights, values, and a knapsack constraint, the input also has a partial order on the items and the requirement that the chosen knapsack be closed under the predecessor relation. This problem is strongly NP-complete ([9]), and the argument can be readily adapted to show that the constrained winner problem is hard likewise.

Theorem 3.1. The constrained winner existence problem for weakly separable voting rules with disjoint tribes is NP-complete, even for SNTV.

For Bloc, it is NP-complete even with a constant number of voters.
Proof. Let \((G, k)\) be an instance of clique. Define an SNTV election with a candidate for every vertex and every edge. For every edge candidate \(e_i\), define a voter that ranks \(e_i\) first and the rest arbitrarily. The set of tribes is the set of singletons. For every \(e_i = (v_1, v_2) \in E\), add the constraints \(\text{AsMany}\{\{v_1\}, \{e_i\}\}\) and \(\text{AsMany}\{\{v_2\}, \{e_i\}\}\). We claim that there’s a winning committee of size \(k + k(k - 1)/2\) with score \(k(k - 1)/2\), if and only if \(G\) has a clique of size \(k\).

First observe that the requirement that \(e_i = (v_1, v_2)\) is on a committee only if \(v_1\) and \(v_2\) are on the committee establishes that the committee is a subgraph – edges cannot be present without their incident vertices. From this we can establish that no committee of size \(k + k(k - 1)/2\) can have more than \(k(k - 1)/2\) points, as that would represent a graph with at least \(k(k - 1)/2 + 1\) edges and at most \(k - 1\) vertices.

In order to have \(k(k - 1)/2\) points, then, we need to have \(k\) vertices and \(k(k - 1)/2\) edges, and this can only be a complete graph of order \(k\), that is to say a clique.

For Bloc, we first claim that clique remains hard if we restrict ourselves to the case where a clique of size \(k\) contains at least half the edges of the graph, i.e. \(2\binom{k}{2} \geq \binom{|V|}{2}\). To see that this is the case, given an instance of clique \((G, k)\) expand \(G\) into \(G'\) by adding \(3|V|\) new vertices, adjacent to each other and to every vertex in \(|V|\). Clearly, \(G'\) contains a clique of size \(k + 3|V|\) if and only if \(G\) contains a clique of size \(k\), and one can verify that \(2\binom{k + 3|V|}{2} \geq \binom{4|V|}{2}\) for \(k \geq 2\).

Consider then an instance of clique with \(2\binom{k}{2} \geq \binom{|V|}{2}\). Define a candidate for every vertex, a candidate for every edge, and \(k + k(k - 1)/2\) dummy candidates. Define one voter that ranks \(k(k - 1)/2\) edges in the top positions in any order, then the dummy candidates, then the other candidates. The second voter will rank the remaining edges first, then the dummy candidates, then the other candidates. Add the constraint \(\text{Exactly}(0, D)\) for the tribe of dummy candidates \(D\), and \(\text{AsMany}\{\{v_1\}, e\}, \text{AsMany}\{\{v_2\}, e\}\) for every edge \(e = (v_1, v_2)\). From hereon replicate the argument for SNTV.

We include the argument for Bloc because Bloc also belongs to a class of rules known as top-\(k\) counting ([7]), and while such rules are hard to solve in general, a wide class of them are fixed-parameter tractable with respect to the number of voters. Now we see that for the constrained problem, this is no longer the case.

The fact that partial-order knapsack is hard, of course, is not surprising since knapsack by itself is already NP-complete. What is key here is that while knapsack is solvable in pseudopolynomial time, i.e. can be solved in polynomial time if all the weights are polynomial in the size of the input, the proof above establishes hardness even if all the weights are zero or one.

However, [9] showed the existence of a pseudopolynomial time solution for partial-order knapsack if the partial order is a forest. If we assume a similar restriction in the constrained winner problem, we will be able to construct partial-order knapsack where the weights are polynomial in the size of the constrained winner instance, and this will show the existence of a polynomial time solution for the constrained winner problem.

**Definition 3.1.** The max exact partial-order knapsack problem takes as input a set of items, \((v_1, w_1), \ldots, (v_n, w_n)\), a partial order on the items \(\succeq\), and a knapsack constraint \(W\).
The output is a subset $X$ of items closed under predecessor with respect to $\geq$, with $\sum_{(v_i,w_i) \in X} w_i = W$ and maximal $\sum_{(v_i,w_i) \in X} v_i$.

**Proposition 3.1** ([9]). If there do not exist distinct $x, y, z$ such that $x \geq y, z \geq y,$ and $x$ is incomparable with $z$, then exact partial-order knapsack is solvable in time polynomial in $n, W$.

**Theorem 3.2.** Let $\geq$ represent the relation induced by the AsMany constraints, and $E_i$ denote the equivalence class of $T_i$ with respect to $\geq$. A set of constraints is said to be sylvan if there do not exist distinct $E_i, E_j, E_r$ for which $E_i \geq E_j$, $E_r \geq E_j$, and $E_i$ is incomparable with $E_r$.

The constrained winner selection problem for a weakly separable voting rule with disjoint tribes and sylvan constraints is solvable in polynomial time.

**Proof.** Construct a directed graph $G$ with vertices $T$ and arcs AsMany($T_i, T_j$) $\in \mathcal{E}$. Take the transitive closure of $G$. In every $T_i$, order the candidates in terms of their score, breaking ties arbitrarily. We say that the $j$th candidate in this ordering has rank $j$. Sort $G$ topologically, and starting with the topologically first $i$ consider every pair $(T_i, T_j) \in G$. If $|T_j| > |T_i|$, delete the $|T_j| − |T_i|$ lowest rank elements of $T_j$.

For every AtMost($p, T_i$) $\in \mathcal{E}$ delete the $|T_j| − p$ lowest rank candidates from all $T_j$ for which $(T_i, T_j) \in \mathcal{E}$.

Initialise a function $r : \mathcal{T} \rightarrow \mathbb{N}$ to $r(T_i) = 0$. For every AtLeast($p, T_i$) $\in \mathcal{E}$, update $r(T_j) = \max(p, r(T_j))$ for all $T_j$ for which $(T_j, T_i) \in G$. Remove the $r(T_i)$ highest rank candidates from all $T_i$ and put them into a set $Y$. Note that in removing these candidates we do not change the rank of the remaining candidates.

Collapse every clique into a single vertex. Where a clique $T_1, \ldots, T_p$ is collapsed in such a way, populate this clique with the $p$-tuples $\{(c_1, \ldots, c_p) \mid c_i \in T_i, \text{rank}(c_1) = \cdots = \text{rank}(c_p)\}$. The rank of a $p$-tuple is the rank of its elements.

For every singleton $c$ in the graph, create an item with weight 1 and value equal to the score of $c$. For every $p$-tuple, create an item with weight $p$ and value equal to the sum of the scores in the $p$-tuple. Define a partial order on the items by setting $x \geq y$ if and only if $x \in T_i, y \in T_j, (T_i, T_j) \in \mathcal{E}$, and $\text{rank}(x) = \text{rank}(y)$. To complete the exact partial-order knapsack instance, the knapsack capacity will be $k − |Y|$.

The above construction is polynomial time. The transitive closure can be found in polynomial time, e.g. with the Floyd–Warshall algorithm, clique detection in a transitive graph reduces to cycle detection, and the other operations are clearly polynomial. The end result is a partial-order knapsack instance where the largest weight of an item is bounded above by the largest clique size, or $|T|$, which is polynomial in the size of the input. Thus this instance can be solved in polynomial time. It remains to show how the solution gives us an election winner.

Recall that $Y$ is the set of candidates removed for the AtLeast constraints. Let $\beta = \text{score}_E(Y)$. We will now show that if $X$ is a solution to the knapsack instance with value $\alpha$ then $X \cup Y$ is a $k$-committee satisfying $\mathcal{C}$ with a score of $\alpha + \beta$, and that if $X$ is a $k$-committee satisfying $\mathcal{C}$ with a score of $\alpha + \beta$ then there exists a solution $X'$ to the knapsack instance with value at least $\alpha$. This will establish that an optimal
solution to the knapsack instance can be used to obtain a constrained election winner by adding the candidates in $Y$.

Suppose that $X$ is a solution to the knapsack instance with value $\alpha$. Observe that every element in $G$ had weight equal to the number of candidates it represented, so $X \cup Y$ is indeed a $k$-committee. It is clear that the score of $X \cup Y$ is $\alpha + \beta$, it remains to show that $X \cup Y$ satisfies $\mathcal{C}$. Constraints of the form $\text{AtMost}(p, T_i)$ are satisfied because we removed all but $p$ candidates from all such $T_i$. The $\text{AtLeast}$ constraints are satisfied by virtue of the candidates in $Y$. For the $\text{AsMany}(T_i, T_j)$ constraints, observe that if $x$ is in the knapsack, then so is its entire $\geq$-predecessor chain until the initial element, $x_1$. If $x_1$ belongs to a vertex in $G$ with zero in-degree, then this chain satisfies all the constraints imposed by the arcs on the graph. If on the other hand $x_1$ belongs to vertex $A$, and $(B, A) \in G$, then it must be the case that the vertex with the same rank as $x_1$, $x_0$, has been removed from $B$. Since $x_0$ will be added to the committee from $Y$, it will still serve to satisfy the constraints.

Suppose that $X$ is a $k$-committee satisfying $\mathcal{C}$ with a score of $\alpha + \beta$. Recall that $r : T \to \mathbb{N}$ is the function telling us how many members from $T_i$ are removed and put into $Y$. Since $X$ satisfies $\mathcal{C}$, it follows that $X$ must contain at least $r(T_i)$ members of each tribe $T_i$. For every $T_i$, remove the $r(T_i)$ highest rank elements of $T_i$ from $X$. By construction, the combined score of all these elements is at most $\beta$, and so the resulting rump committee $X'$ has score $\alpha' \geq \alpha$.

To obtain a knapsack solution from $X'$, simply take the $|X' \cap T_i|$ highest rank candidates from all $T_i$, and this solution has value at least $\alpha$ as required.

\section{Small number of tribes and fixed-parameter tractability}

If the number of tribes is slow we can obtain a polynomial time solution via mixed integer linear programming, using similar techniques to the result for top-$k$ counting rules by [7].

**Theorem 4.1.** The constrained winner election problem for weakly separable rules is fixed parameter tractable with respect to the number of tribes.

**Proof.** Intuitively, if the number of tribes is constant this problem can be solved by brute force. Observe that if $X$ is a committee satisfying constraints $\mathcal{C}$, then if $x \in X$, $y \notin Y$, $\text{score}_E(y) > \text{score}_E(x)$ and for all tribes $T_i$, $x \in T_i$ if and only if $y \in T_i$, then $Y = (X \setminus \{x\}) \cup \{y\}$ is also a committee satisfying the constraints, and $\text{score}_E(Y) \geq \text{score}_E(X)$. In other words if two candidates have the same tribal affiliations, we will never violate a constraint by swapping a low scoring one for a high scoring one, and doing so will only increase the score.

Since a constant $p$ sets gives rise to a Venn diagram with a constant $q$ regions, we need only consider all ways of choosing the best $k$ candidates from every region, of which there are at most $k^q$.

To obtain a fixed parameter tractable algorithm, we will recast the intuition above as a mixed integer linear program. Such a program is fixed parameter tractable in $p$ if the number of integer variables is a function of $p$ alone ([12]).
Let $D_1, \ldots, D_q$ enumerate the regions of the Venn diagram induced by $T$. Note that $q \leq 2^p$. Introduce the integer variables $d_1, \ldots, d_q$, the interpretation of $d_i$ being the number of elements taken from $D_i$. Introduce the real indicator variables $(s_{i,j})_{i \leq q, j \leq k}$ with the interpretation that $s_{i,j} = 1$ if and only if at least $j$ elements are taken from $D_i$, and the auxiliary real variables $(a_{i,j})_{i,j \leq p}$ to handle the AsMany constraints.

The constant values $c_{i,j}$ will represent the score of the $j$th highest scoring candidate from $D_i$. The resulting system is in fig. 1.

Constraint 1 ensures the committee is of size $k$, constraints 2 through 4 ensure the satisfaction of $C$, and constraint 5 establishes the relation between the $d_i$ variables and the objective function. Clearly if the system arrives at a solution where all $s_{i,j}$ are integral, we have found a maximal score $k$-committee.

Suppose then that the solution is such that there exists a $0 < s_{i,j} < 1$. Since $\sum_{j \leq k} s_{i,j} = d_i$, and $d_i$ is an integer, it follows there exists another $0 < s_{i,j'} < 1$. Without loss of generality, let $j' > j$. Since $c_{i,j} \geq c_{i,j'}$ by definition, the value of the solution will not decrease if we transfer the weight from $s_{i,j'}$ to $s_{i,j}$, after which we will have reduced the number of non-integral values by one. By repeating this process for each non-integral $s_{i,j}$, we will find an integral solution with the same value. 

$$
\max \sum_{i \leq q} \sum_{j \leq k} s_{i,j} c_{i,j}
$$

s.t.

1) $\sum_{i \leq q} d_i = k$,

2) $\sum_{D_j \subseteq T_i} d_j \geq r$, for all $\text{AtLeast}(r, T_i)$

3) $\sum_{D_j \subseteq T_i} d_j \leq r$, for all $\text{AtMost}(r, T_i)$

4) $\sum_{D_j \subseteq T_i} d_j \geq a_{i,r}$, $\sum_{D_j \subseteq T_i} d_j \leq a_{i,r}$, for all $\text{AsMany}(T_i, T_r)$

5) $\sum_{i \leq q} \sum_{j \leq k} s_{i,j} = d_i$,

6) $0 \leq s_{i,j} \leq 1$, $i \leq q, j \leq k$

Figure 1: A mixed integer formulation of the constrained winner problem. Integer variables in typewriter font.
5 Discussion

We have introduced a framework for specifying an election with constraints, and have identified two restrictions that make the problem tractable for separable voting rules: tree-like dependencies, and a small number of tribes. For more general constraints, we have shown that the problem is NP-hard, even if we restrict ourselves to the case of disjoint tribes and the SNTV voting rule.

Since our focus in this work was on exact algorithms we did not consider other classes of committee scoring rules such as top-k counting, representation focused, or ordered weighted average rules, as even the unconstrained election problem for these rules is hard. Future work could involve developing approximation algorithms to find approximately optimal constrained election winners for the hard rules, or indeed approximately optimal committees for separable rules, in the setting where we do not restrict ourselves to easy constraints.

We have argued that for the type of restrictions seen in practice we do not need a rich language, but simply a list of constraints specifying how many candidates we should take from each tribe, and how one tribe’s representation should relate to another’s. While a sensible assumption for elections involving humans, when dealing with artificial agents one might be tempted to allow more general restrictions. However, it is not clear whether we will gain anything by this route – we can already express propositional relations by using general (i.e. not disjoint) tribes, so adding a layer of logic on top of this may be superfluous.

Perhaps more interesting is the question is how to elect a committee with constraints for rules that do not produce a ranking of committees, e.g. the Condorcet based rules ([8, 1, 5]). This will require an axiomatic approach to consider what exactly we want from such a committee – obviously we throw away stability once we have reserved seats, but what can we implement in its place?

References

[1] Salvador Barberà and Danilo Coelho. How to choose a non-controversial list with k names. Social Choice and Welfare, 31(1):79–96, Jun 2008.

[2] Nadja Betzler, Arkadii Slinko, and Johannes Uhlmann. On the computation of fully proportional representation. Journal of Artificial Intelligence Research, 47:475–519, 2013.

[3] John Duggan and Thomas Schwartz. Strategic manipulability is inescapable: Gibbard-satterthwaite without resoluteness. Working Papers 817, California Institute of Technology, Division of the Humanities and Social Sciences, 1992.

[4] Edith Elkind, Piotr Faliszewski, Piotr Skowron, and Arkadii Slinko. Properties of multiwinner voting rules. Social Choice and Welfare, 48(3):599–632, Mar 2017.

[5] Edith Elkind, Jérôme Lang, and Abdallah Saffidine. Condorcet winning sets. Social Choice and Welfare, 44(3):493–517, Mar 2015.
[6] Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. Committee scoring rules: Axiomatic classification and hierarchy. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI'16, pages 250–256. AAAI Press, 2016.

[7] Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. Multi-winner analogues of the plurality rule: axiomatic and algorithmic perspectives. Social Choice and Welfare, 51(3):513–550, Oct 2018.

[8] Peter C Fishburn. Majority committees. Journal of Economic Theory, 25(2):255–268, 1981.

[9] D. S. Johnson and K. A. Niemi. On knapsacks, partitions, and a new dynamic programming technique for trees. Mathematics of Operations Research, 8(1):1–14, 1983.

[10] D. Marc Kilgour. Approval balloting for multi-winner elections. In Jean-François Laslier and M. Remzi Sanver, editors, Handbook on Approval Voting, pages 105–124. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[11] D. Marc Kilgour and Erica Marshall. Approval balloting for fixed-size committees. In Dan S. Felsenthal and Moshé Machover, editors, Electoral Systems: Paradoxes, Assumptions, and Procedures, pages 305–326. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[12] H. W. Lenstra. Integer programming with a fixed number of variables. Mathematics of Operations Research, 8(4):538–548, 1983.

[13] Tyler Lu and Craig Boutilier. Budgeted social choice: From consensus to personalized decision making. In Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume One, IJCAI’11, pages 280–286. AAAI Press, 2011.

[14] Ariel D. Procaccia, Jeffrey S. Rosenschein, and Aviv Zohar. On the complexity of achieving proportional representation. Social Choice and Welfare, 30(3):353–362, Apr 2008.

[15] Piotr Skowron, Piotr Faliszewski, and Jerome Lang. Finding a collective set of items: From proportional multirepresentation to group recommendation. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, AAAI’15, pages 2131–2137. AAAI Press, 2015.

[16] Alan D. Taylor. The manipulability of voting systems. The American Mathematical Monthly, 109(4):321–337, 2002.