Electing a committee with dominance constraints

Egor Ianovski*

International Laboratory of Game Theory and Decision Making, National Research University Higher School of Economics, St Petersburg

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 candidates are given labels as an abstraction of gender, religion, ethnicity, and other attributes, and the election outcome is constrained by interval constraints – of the form "Between 3 and 5 candidates with label X" – and dominance constraints – "At least as many candidates with label X as with label Y". 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 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

* Support from the Basic Research Program of the National Research University Higher School of Economics is gratefully acknowledged.

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 [5], 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 [16, 2, 15].

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. Ordered weighted average operators where introduced by [9], which led to the superclass of ordered weighted average rules [18], 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, and has been present in the approval-voting literature [11, 12]. The specific question of how to optimise the score of a committee subject to range constraints was studied by [3]. They present approximation algorithms for submodular scoring functions, while for separable scores they present exact solutions for problems with simple label structures and hardness results otherwise. Independently, [4] introduce a very similar model where they present approximation algorithms for submodular scores. The work of [1] considers the question of how to elect a committee if the constraints cannot be satisfied, and presents an algorithm for finding an ordinally-optimal committee that comes the closest to satisfying the constraints. Also related is the work on apportionment of [13] which considers how to apportion seats with an arbitrary number of diversity constraints.

1.2 Our contribution

We extend the models of [3] and [4] by introducing dominance constraints, and show that the constrained election problem with these constraints for disjoint labels is solvable in polynomial time if the dominance relation is tree-like and NP-hard in the general case. For arbitrary label structures, we show that the problem is fixed-parameter tractable in the number of labels.

2 Preliminaries

2.1 Committee scoring rules

Let $[m]$ be the set of integers $\{1, \dots, m\}$, and $[m]_k$ be the set of all length-k increasing sequences of numbers from $[m]$. Given two sequences, $I = (i_1, ..., i_k)$ and $J = (j_1, ..., j_k)$, 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 $\mathrm{pos}_{v_i}(c), c \in C$, 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 $\mathrm{pos}_{v_i}(X)$ to denote the sequence that we obtain by sorting the set $\{\mathrm{pos}_{v_i}(c) \mid c \in X\}$ in increasing order. Naturally, $\mathrm{pos}_{v_i}(c) \in [m]$ and $\mathrm{pos}_{v_i}(X) \in [m]_k$.

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 $\mathrm{pos}_{v_i}(X)$.

Definition 1. *A* committee scoring function *for m candidates and committee size k is a function $f_{m,k} : [m]_k \to \mathbb{R}^+$ 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 $\mathcal{R}_f$ that given an election E and an integer k outputs all k-committees that maximise $score_E(X) = \sum_{v_i \in V} f_{m,k}(\mathrm{pos}_{v_i}(X))$.*

[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.

[1] Subject to assumptions about the underlying scoring functions being polynomial-time computable

Definition 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* $\gamma_{m,k} : [m] \to \mathbb{R}^+$ *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 $\mathcal{R}_f$ *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:

$$\begin{aligned}\text{score}_E(X) &= \sum_{v_i\in V} f_{m,k}(\text{pos}_{v_i}(X)) \\ &= \sum_{v_i\in V}\sum_{j\in X} \gamma_{m,k}(\text{pos}_{v_i}(j)) \\ &= \sum_{j\in X}\sum_{v_i\in V} \gamma_{m,k}(\text{pos}_{v_i}(j)).\end{aligned}$$

We will thus refer to $\sum_{v_i\in V} \gamma_{m,k}(\text{pos}_{v_i}(j))$ as the score of j, or $\text{score}_E(j)$.

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, φ_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 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. The constraints used in practice tend to be very simple – allocating a number of seats (e.g., 26 bishops in the House of Lords), setting a lower bound (at least one Pleb consul), or establishing parity between groups (as many Democrats as Republicans). Most of these can be captured through the use of *interval constraints*

[3][4], i.e. setting a numerical lower and upper bound on the number of candidates of a certain type. In this paper we also introduce *dominance constraints*, which will allow us to require that one group has at least as many candidates as another, without recourse to exact numbers.

This gives us our constraints:

- Interval(p, q, X) : between p and q members of the committee from X.
- Dominance(X, Y) : at least as many members of the committee are from X as from Y.

We can now define the algorithmic problems of interest.

Definition 3. *The* constrained winner election *problem for a committee scoring rule $\mathcal{R}$ is the problem that takes an election E, a set of constraints $\mathfrak{C}$, and a set of* labels $\Lambda = \lambda_1, \ldots, \lambda_p \subseteq C$ *as input. The output is some k-committee X that maximises the score out of all the committees that satisfy $\mathfrak{C}$.*

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

Unfortunately, if Λ is arbitrary then simply determining whether there exists a committee satisfying $\mathfrak{C}$ is NP-hard.

Proposition 1. *It is* NP*-hard to determine whether there exists a committee X satisfying a set of constraints $\mathfrak{C}$. Thus the constrained winner existence problem is* NP*-complete for any onto committee selection rule.*

Proof. Hardness for interval constraints has been established by [3] and [4]. For dominance constraints, we reduce from vertex cover. Construct an election with a candidate for every vertex in the graph, $c_1, \ldots, c_m$. Define a label for every edge in the graph, consisting of the vertices incident on the edge, and a label for every singleton candidate. For every edge e, introduce the constraints Dominance$(e, \{c_1\}), \ldots,$ Dominance$(e, \{c_m\})$. Since at least one candidate must be elected to form a size k committee, the constraints establish that at least one vertex from every edge must be chosen.

We will thus need to impose some structure on the admissible lists of labels to proceed. We consider two natural cases: where the labels are disjoint (corresponding to the 1-layered case of [3], and the $\Delta = 1$ case of [4]), and where there is only a small number of labels.

3 Disjoint labels and knapsack

We note that the dominance constraints impose an ordering on the labels – if we imagine we are building the committee one candidate at a time, and X dominates Y, then we must take a candidate from X before we take one from Y. If there are several labels 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. In our case the relation is a preorder, since it is possible for X and Y to dominate each other (thus requring that the committee has the same number of candidates from X and Y), but, since a preorder generalises a partial order, the strong NP-completeness of partial order knapsack ([10]) carries over to our case as well.

Theorem 1. *The constrained winner existence problem for weakly separable voting rules with disjoint labels 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, $G = (\mathcal{V}, \mathcal{E})$. 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 labels is the set of singletons. For every $e_i = (v_1, v_2) \in \mathcal{E}$, add the constraints $\text{Dominance}(\{v_1\}, \{e_i\})$ and $\text{Dominance}(\{v_2\}, \{e_i\})$. 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. $2\binom{k}{2} \geq \binom{|\mathcal{V}|}{2}$. To see that this is the case, given an instance of clique (G, k) expand G into G' by adding $3|\mathcal{V}|$ new vertices, adjacent to each other and to every vertex in $|\mathcal{V}|$. Clearly, G' contains a clique of size $k + 3|\mathcal{V}|$ if and only if G contains a clique of size k, and one can verify that $2\binom{k+3|\mathcal{V}|}{2} \geq \binom{4|\mathcal{V}|}{2}$ for $k \geq 2$.

Consider then an instance of clique with $2\binom{k}{2} \geq \binom{|\mathcal{V}|}{2}$. 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 $\text{Interval}(0, 0, D)$ for the label of dummy candidates D, and $\text{Dominance}(\{v_1\}, e)$, $\text{Dominance}(\{v_2\}, e)$ 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. Here we see that the constrained problem is hard even with a constant number.

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, [10] 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 4. *The* 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 $\succeq$, with $\sum_{(v_i,w_i)\in X} w_i = W$ and maximal $\sum_{(v_i,w_i)\in X} v_i$.*

Proposition 2 ([10]). *If there do not exist distinct x, y, z such that $x \succeq y$, $z \succeq y$, and x is incomparable with z, then exact partial-order knapsack is solvable in time polynomial in n, W.*[2]

In the case of constrained elections we can generalise this slightly as we do not require the preordering of labels to be a strict forest, but a forest modulo cliques – equivalently, the partial order on equivalence classes is a forest. Cases like $\lambda_1 \sim \lambda_2 \succeq \lambda_3$ will not break the result.

Theorem 2. *Let $\succeq$ represent the preorder induced by the* Dominance *constraints, and E_i denote the equivalence class of λ_i with respect to $\succeq$. A set of constraints is said to be* sylvan *if there do not exist distinct E_i, E_j, E_r for which $E_i \succeq E_j$, $E_r \succeq E_j$, and E_i is incomparable with E_r.*

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

Proof. Construct a directed graph G with vertices Λ and arcs Dominance$(\lambda_i, \lambda_j) \in \mathfrak{C}$. Take the transitive closure of G – note that the arcs of the resulting graph are precisely the preorder $\succeq$.

In every λ_i, order the candidates in terms of their score, breaking ties arbitrarily. We say that the jth candidate in this ordering has rank j. Sort G topologically, and starting with the topologically first i consider every pair $(\lambda_i, \lambda_j) \in G$. If $|\lambda_j| > |\lambda_i|$, delete the $|\lambda_j| - |\lambda_i|$ lowest rank elements of λ_j.

For every Interval$(p, q, \lambda_i \in \mathfrak{C}$, first delete the $|\lambda_j| - q$ lowest rank candidates from all λ_j for which $(\lambda_i, \lambda_j) \in G$.

[2] [10] do not consider the exact variant, but the dynamic programming solution can be modified to handle it in the standard way.

Second, initialise a function $r : \Lambda \rightarrow \mathbb{N}$ to $r(\lambda_i) = 0$. For every λ_j for which $(\lambda_j, \lambda_i) \in G$, update $r(\lambda_j) = \max(p, r(\lambda_j))$. Remove the $r(\lambda_i)$ highest rank candidates from all λ_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 $\lambda_1, \ldots, \lambda_p$ is collapsed in such a way, populate this clique with the p-tuples $\{(c_1, \ldots, c_p) \mid c_i \in \lambda_i,\ \text{rank}(c_1) = \cdots = \text{rank}(c_p)\}$. 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 $x \in \lambda_i$, $y \in \lambda_j$, $(\lambda_i, \lambda_j) \in \mathcal{E}$, and $\text{rank}(x) = \text{rank}(y)$. To complete the exact partial-order knapsack instance, the knapsack capacity will be k.

The above construction is polynomial time. The transitive closure can be found in polynomial time, e.g. with the FloydWarshall 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 $|\Lambda|$, 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 lower bounds of the interval constraints. Let $\beta = \text{score}_E(Y)$. We will now show that if X is a solution to the knapsack instance with value α then $X \cup Y$ is a k-committee satisfying $\mathfrak{C}$ with a score of $\alpha + \beta$, and that if X is a k-committee satisfying $\mathfrak{C}$ with a score of $\alpha + \beta$ then there exists a solution X' to the knapsack instance with value at least α. 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 α. 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 $\mathfrak{C}$. The upper bound of $\text{Interval}(p, q, \lambda_i)$ is satisfied because we removed all but q candidates from all such λ_i. The lower bounds are satisfied by virtue of the candidates in Y. For the $\text{Dominance}(\lambda_i, \lambda_j)$ 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 $\mathfrak{C}$ with a score of $\alpha + \beta$. Recall that $r : \Lambda \rightarrow \mathbb{N}$ is the function telling us how many members from λ_i are removed and put into Y. Since X satisfies $\mathfrak{C}$, it follows that X must contain at least $r(\lambda_i)$

members of each label λ_i. For every λ_i, remove the $r(\lambda_i)$ highest rank elements of λ_i from X. By construction, the combined score of all these elements is at most β, and so the resulting rump committee X' has score $\alpha' \geq \alpha$.

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

4 Small number of labels and fixed-parameter tractability

If the number of label 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] and Theorem 10 of [3].

Theorem 3. *The constrained winner election problem for weakly separable rules is fixed parameter tractable with respect to the number of labels.*

Proof. Intuitively, if the number of labels is constant this problem can be solved by brute force. Observe that if X is a committee satisfying constraints $\mathfrak{C}$, then if $x \in X$, $y \notin Y$, $\text{score}_E(y) > \text{score}_E(x)$ and for all labels λ_i, $x \in \lambda_i$ if and only if $y \in \lambda_i$, 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 labels, 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 ([14]).

Let $D_1, \ldots, D_q$ enumerate the regions of the Venn diagram induced by Λ. Note that $q \leq 2^p$. Introduce the integer variables $\mathtt{d_1}, \ldots, \mathtt{d_q}$, the interpretation of $\mathtt{d_i}$ 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..

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

Constraint 1 ensures the committee is of size k, constraints 2 ans 3 ensure the satisfaction of $\mathfrak{C}$, and constraint 4 establishes the relation between the $\mathtt{d_i}$ 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 $\sum_{j \leq k} s_{i,j} = \mathtt{d_i}$, and $\mathtt{d_i}$ 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.

$$\begin{aligned}
\max \quad & \sum_{i \le q} \sum_{j \le k} s_{i,j} c_{i,j} \\
\text{s.t.} \\
1) \quad & \sum_{i \le q} \mathtt{d_i} = k, \\
2) \quad & \sum_{D_j \subseteq \lambda_i} \mathtt{d_j} \ge r, \\
& \sum_{D_j \subseteq \lambda_i} \mathtt{d_j} \le t, && \text{for all Interval}(r, t, \lambda_i) \\
3) \quad & \sum_{D_j \subseteq \lambda_i} \mathtt{d_j} \ge \sum_{D_j \subseteq \lambda_t} \mathtt{d_j}, && \text{for all Dominance}(\lambda_i, \lambda_r) \\
4) \quad & \sum_{i \le q} \sum_{j \le k} s_{i,j} = \mathtt{d_i}, \\
5) \quad & 0 \le s_{i,j} \le 1, && i \le q, j \le k
\end{aligned}$$

Fig. 1. A mixed integer formulation of the constrained winner problem. Integer variables in typewriter font.

5 Future directions

The work of [3], [4], and the present work all reduce to optimising a certain objective function over the space of viable committees. This model is both amenable to an algorithmic approach and has a ready economic interpretation in terms of utility maximisation, but is at odds with the standard model of voting in social choice theory which is purely ordinal. In this paper we have interpreted this approach as optimising the ordinal ranking that is the outcome of the cardinal scores of a scoring rule, but scoring rules are just one of many voting procedures that can be used to elect a committee.

A purely ordinal model is used by [1], but it is a very specific one – we are given a global preference order over the candidate and have to choose an optimal committee accordingly. This could be interpreted as optimising a best-k rule in the sense of [5]. It would be interesting to see what could be done for more general varieties of committee selection rules, such as the Condorcet-based approaches of [8] and [17].

References

1. Aziz, H.: A rule for committee selection with soft diversity constraints. CoRR **abs/1803.11437** (2018), http://arxiv.org/abs/1803.11437
2. Betzler, N., Slinko, A., Uhlmann, J.: On the computation of fully proportional representation. Journal of Artificial Intelligence Research **47**, 475–519 (2013)

3. Bredereck, R., Faliszewski, P., Igarashi, A., Lackner, M., Skowron, P.: Multiwinner elections with diversity constraints. In: Thirty-Second AAAI Conference on Artificial Intelligence (2018)
4. Celis, L.E., Huang, L., Vishnoi, N.K.: Multiwinner voting with fairness constraints. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence. pp. 144–151. AAAI Press (2018)
5. Elkind, E., Faliszewski, P., Skowron, P., Slinko, A.: Properties of multiwinner voting rules. Social Choice and Welfare **48**(3), 599–632 (Mar 2017). https://doi.org/10.1007/s00355-017-1026-z, https://doi.org/10.1007/s00355-017-1026-z
6. Faliszewski, P., Skowron, P., Slinko, A., Talmon, N.: Committee scoring rules: Axiomatic classification and hierarchy. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. pp. 250–256. IJCAI'16, AAAI Press (2016), http://dl.acm.org/citation.cfm?id=3060621.3060657
7. Faliszewski, P., Skowron, P., Slinko, A., Talmon, N.: Multiwinner analogues of the plurality rule: axiomatic and algorithmic perspectives. Social Choice and Welfare **51**(3), 513–550 (Oct 2018). https://doi.org/10.1007/s00355-018-1126-4, https://doi.org/10.1007/s00355-018-1126-4
8. Fishburn, P.C.: Majority committees. Journal of Economic Theory **25**(2), 255 – 268 (1981). https://doi.org/https://doi.org/10.1016/0022-0531(81)90005-3, http://www.sciencedirect.com/science/article/pii/0022053181900053
9. Goldsmith, J., Lang, J., Mattei, N., Perny, P.: Voting with rank dependent scoring rules. In: Twenty-Eighth AAAI Conference on Artificial Intelligence (2014)
10. Johnson, D.S., Niemi, K.A.: On knapsacks, partitions, and a new dynamic programming technique for trees. Mathematics of Operations Research **8**(1), 1–14 (1983), http://www.jstor.org/stable/3689406
11. Kilgour, D.M.: Approval balloting for multi-winner elections. In: Laslier, J.F., Sanver, M.R. (eds.) Handbook on Approval Voting, pp. 105–124. Springer Berlin Heidelberg, Berlin, Heidelberg (2010). https://doi.org/10.1007/978-3-642-02839-7_6, https://doi.org/10.1007/978-3-642-02839-7_6
12. Kilgour, D.M., Marshall, E.: Approval balloting for fixed-size committees. In: Felsenthal, D.S., Machover, M. (eds.) Electoral Systems: Paradoxes, Assumptions, and Procedures, pp. 305–326. Springer Berlin Heidelberg, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-20441-8_12, https://doi.org/10.1007/978-3-642-20441-8_12
13. Lang, J., Skowron, P.: Multi-attribute proportional representation. Artificial Intelligence **263**, 74–106 (2018)
14. Lenstra, H.W.: Integer programming with a fixed number of variables. Mathematics of Operations Research **8**(4), 538–548 (1983). https://doi.org/10.1287/moor.8.4.538, https://doi.org/10.1287/moor.8.4.538
15. Lu, T., Boutilier, C.: Budgeted social choice: From consensus to personalized decision making. In: Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume One. pp. 280–286. IJCAI'11, AAAI Press (2011). https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-057, http://dx.doi.org/10.5591/978-1-57735-516-8/IJCAI11-057
16. Procaccia, A.D., Rosenschein, J.S., Zohar, A.: On the complexity of achieving proportional representation. Social Choice and Welfare **30**(3), 353–362 (Apr 2008). https://doi.org/10.1007/s00355-007-0235-2, https://doi.org/10.1007/s00355-007-0235-2

17. Ratliff, T.C.: Some startling inconsistencies when electing committees. Social Choice and Welfare **21**(3), 433–454 (Dec 2003). https://doi.org/10.1007/s00355-003-0209-y, https://doi.org/10.1007/s00355-003-0209-y
18. Skowron, P., Faliszewski, P., Lang, J.: Finding a collective set of items: From proportional multirepresentation to group recommendation. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. pp. 2131–2137. AAAI'15, AAAI Press (2015), http://dl.acm.org/citation.cfm?id=2886521.2886617