Compilation Complexity of Multi-Winner Voting Rules (Student Abstract)

Neel Karia,[1] **Jérôme Lang,** [2]
[1] Department of Computer Science and Engineering, IIT Kharagpur
[2] CNRS, LAMSADE, PSL, Université Paris-Dauphine
karianeel@iitkgp.ac.in, lang@lamsade.dauphine.fr

Abstract

Compiling the votes of a subelectorate consists of storing the votes of a subset of voters in a compressed form, such that the winners can still be determined when additional votes are included. This leads to the notion of compilation complexity, which has already been investigated for single-winner voting rules. We perform a compilation complexity analysis of several common multi-winner voting rules.

Introduction

Voting is a common way by which a group of agents make joint decisions. However, in many contexts, the votes are not obtained at the same time or at the same place. In such scenarios, we might want to preprocess the information contained in the votes that are already available. Compiling a set of votes means storing the information contained in these votes using as little space as possible in such a way that when the rest of the votes are known, the winner(s) can be determined. The compilation complexity of a voting rule is the worst-case size of the most succinct compilation. Compilation has two advantages: first, the votes of the subelectorate can be stored succinctly; second, the storage of the information contained in the ballots of the subelectorate can be done in an anonymous, yet verifiable manner. Compilation of single-winner rules has been studied. Here we initiate the compilation of multi-winner voting rules.

Multi-Winner Voting Rules

Let A be a set of candidates, with $|A| = m$. Let $[m] = \{1, \ldots, m\}$. Let $\mathcal{P}_A$ be a set of *votes*, which depending on the rule used is either the set of all linear orders over A (ranked ballots) or the set of subsets of A (approval ballots). A *(partial or complete) profile* P is a collection $(V_1, \ldots, V_n)$ of n votes, for some n; it is therefore a member of $\cup_{n \geq 1} \mathcal{P}_A^n$.

Let $\mathcal{S}_k(A)$ be the set of all subsets of A of size k, called *(k-)committees*. A (resolute) multi-winner voting rule is a function f that maps any profile P and any $k \in [m]$ to a k-committee $f(P, k) \in \mathcal{S}_k(A)$.

We present below the multi-winner rules, further described in (Faliszewski et al. 2017; Lackner and Skowron 2018) which we will need in the rest of the paper. All of them except one (sequential plurality) are defined via *scores*: given a profile $P = (V_1, \ldots, V_n)$, a score $s(S, P)$ is associated with each committee $S \in \mathcal{S}_k(A)$, and the winning committee is the one that maximizes $s(S, P)$. If P contains a single vote V then we note $s(S, V)$ instead of $s(S, P)$. The multi-winner rule that selects a committee maximizing s is denoted by f_s. In case we have a tie between two or more committees, there is a tie-breaking mechanism (usually a priority relation over committees) that will output a single committee.

We start with these five rules whose input is a profile consisting of ranked ballots. The first four are defined via scores, and for these four, the score of a k-committee is the sum of the scores $s(S, V_i)$ it gets from all votes $V_1, \ldots, V_n$.

- **Single Non Transferable Vote (SNTV)**: $s(S, V_i) = 1$ if S contains the top candidate of V_i; else $s(S, V_i) = 0$.
- **Bloc**: $s(S, V_i)$ is the number of candidates in S ranked in the first k positions of V_i.
- $k-$**Borda**: $s(S, V_i)$ is the sum of the Borda scores of the candidates in S; the Borda score of a candidate ranked in position j in a vote is $m - j$.
- **Chamberlin-Courant** ($\beta-$**CC**): $s(S, V_i)$ is the Borda score w.r.t. V_i of the best candidate in S according to V_i.
- **Sequential Plurality (SeqPlu)**: We proceed in rounds. Initially, $S = \emptyset$. The candidate ranked first by the largest number of votes is added to S, removed from the profile, and the procedure is repeated k times (breaking ties if necessary). The output is S.

SNTV, Bloc, $k-$Borda and $\beta-$CC are all specific cases of the larger family of *committee scoring rules* (Skowron, Faliszewski, and Slinko 2019), for which (1) $s(S, P) = \sum_{V_i \in P} s(S, V_i)$, and (2) $s(S, V_i)$ is a function of the vector containing the ranks of the elements of S in V_i. Moreover, SNTV, Bloc and $k-$Borda are *decomposable* committee scoring rules: there is a score function over candidates $s(x, V)$ such that $s(S, V) = \sum_{x \in S} s(x, V)$.

For the next two rules, the input is a profile consisting of approval ballots.

- **Approval Voting (AV)**: The winning committee consists of the k candidates that are approved most frequently.
- **Approval-based Chamberlin-Courant** ($\alpha-$**CC**): $s(S, P)$ is the number of votes in P that intersect S.

Compilation Functions, Compilation Complexity

Compilation complexity was introduced by (Chevaleyre et al. 2009) for single-winner voting rules. It was further studied in (Xia and Conitzer 2010). Its extension to multi-winner rules is straightforward so we define it directly in the latter context.

Let f be a multi-winner voting rule. Let $k \in [m]$ be fixed. Consider two profiles $P, Q \in \mathcal{P}_A^n$, that contain the votes of subelectorates composed of n voters. We say that P and Q are *f-equivalent*, which we denote by $P \sim_f Q$, if for each $t \geq 0$ and each profile $T \in \mathcal{P}_A^t$, we have $f(P \cup T, k) = f(Q \cup T, k)$. (Clearly, $\sim_f$ is an equivalence relation.)

A function $\sigma : \mathcal{P}_A^n \times [m] \to \{0,1\}^*$ is a *compilation function for f* if there exists a function $\rho : \{0,1\}^* \times \mathcal{P}_A^* \times [m] \to A$ such that for all $P \in \mathcal{P}_A^n$, $t \geq 0$ and $T \in \mathcal{P}_A^t$, we have $\rho(\sigma(P), T, k) = f(P \cup T, k)$. We denote by $size(\sigma)$ the number of bits needed to represent $\sigma(P)$. The *compilation complexity* of f, denoted $C(f)$, is the minimum value of $size(\sigma)$ over all compilation functions for f. A useful lemma (established for single-winner rules and easily generalizable to multi-winner rules) is that $C(f)$ is the logarithm of the number of equivalence classes for $\sim_f$.

Results for Ranking-Based Rules

Let $k \in [m]$ be fixed. It is easy to derive a sufficient condition for all committee scoring rules:

Lemma 1. *Let f_s be a committee scoring rule based on scoring function s. If for all committees $S \in \mathcal{S}_k(A)$ we have $s(S, P) = s(S, Q)$ then $P \sim_{f_s} Q$.*

A weaker sufficient condition holds for decomposable rules. For $P = (V_1, \ldots, V_n)$ let $s(x, P) = \sum_i s(x, V_i)$.

Lemma 2. *Let f'_s be a decomposable committee scoring rule based on score function s. If for every candidate $x \in A$ we have $s(x, P) = s(x, Q)$ then $P \sim_{f'_s} Q$.*

This applies to SNTV, k-Borda and Bloc. Moreover, for these three rules, this condition is also necessary.

Proposition 1. *Let f be SNTV, k-Borda or Bloc and let s be the corresponding score function. Then $P \sim_f Q$ if and only if for every $x \in A$ we have $s(x, P) = s(x, Q)$.*

Finally, for $\beta-$CC, the general sufficient condition for committee scoring rules is necessary.

Proposition 2. *$P \sim_{\beta-CC} Q$ if and only if for all $S \in \mathcal{S}_k(A)$, $s^{\beta-CC}(S, P) = s^{\beta-CC}(S, Q)$.*

As a corollary of Propositions 1 and 2 we get:

Corollary 1.

1. $C(SNTV) = \Theta\left(m \log\left(1 + \frac{n}{m}\right) + n \log\left(1 + \frac{m}{n}\right)\right)$
2. $C(k - Borda) = \Theta(m \log nm)$
3. *Let $\hat{k} = \min(k, m-k)$. Then* $C(Bloc) = \Theta\left(m \log\left(1 + \frac{n\hat{k}}{m}\right) + n\hat{k} \log\left(1 + \frac{m}{n\hat{k}}\right)\right)$
4. $\log(n(m-k)) \leq C(\beta - CC) \leq \binom{m}{k} \log(n(m-k))$

For each vote V let V^j be the top-j truncation of V, where $j \in [k]$. For example, if $V = (abcd)$ then $V^2 = (ab)$. Given a profile P, for each ordered sequence of j candidates θ^j, consider $N(P, \theta^j)$ to be the number of votes V in P such that $V^j = \theta^j$. For instance, if $P = (abcd, abdc, acbd, dabc, dabc)$ and $k = 2$, then $N(P, ab) = N(P, da) = 2$, $N(P, ac) = 1$, and $N(P, \theta^k) = 0$ for $\theta^k \neq ab, ac, da$.

Proposition 3. *$P \sim_{SeqPlu} Q$ if and only if $N(P, \theta^k) = N(Q, \theta^k)$ for each ordered sequence of k candidates θ^k.*

Corollary 2. $C(SeqPlu)$
$= \Theta\left(\frac{m!}{(m-k)!} \log\left(1 + \frac{n(m-k)!}{m!}\right) + n \log\left(1 + \frac{m!}{n(m-k)!}\right)\right)$

Results for Approval-Based Rules

Let $s^{AV}(a, P)$ be the approval score of a with respect to P.

Proposition 4. *$P \sim_{AV} Q$ if and only if for all $a, b \in A$, $s^{AV}(a, P) - s^{AV}(b, P) = s^{AV}(a, Q) - s^{AV}(b, Q)$*

Corollary 3. $C(f_{AV}) = \Theta(m \log n)$

Let $s^{\alpha-CC}(S, P)$ be the α-CC score of a k-committee S with the ballot set P and W be the winning committee.

Proposition 5. *$P \sim_{\alpha-CC} Q$ holds if and only if for all $S \in \mathcal{S}_k(A)$, $s^{\alpha-CC}(W, P) - s^{\alpha-CC}(S, P) = s^{\alpha-CC}(W, Q) - s^{\alpha-CC}(S, Q)$.*

Corollary 4. $C(\alpha - CC) = O\left(\binom{m}{k} \log n\right)$ *and* $C(\alpha - CC) = \Omega\left(\left\lfloor \frac{m}{k} \right\rfloor \log n\right)$

The reason why the lower and upper bounds for $\alpha-$CC and $\beta-$CC do not match is that it is not easy to count the number of functions from $\mathcal{S}_k(A)$ to $\mathbb{N}$ that correspond to $s^{\alpha-CC}(\cdot, P)$ or $s^{\beta-CC}(\cdot, P)$ for some profile P, because of the dependencies between the scores of the different committees. We are working on obtaining better bounds.

Future Work

We have obtained results for Gehrlein-stable rules, $\theta-$winning Sets and Proportional Approval Voting (PAV), which are not presented here due to lack of space. We are also working on Single Transferable Vote and Monroe.

References

Chevaleyre, Y.; Lang, J.; Maudet, N.; and Ravilly-Abadie, G. 2009. Compiling the Votes of a Subelectorate. In *Proceedings of the 21st International Joint Conference on Artifical intelligence*, 97–102.

Faliszewski, P.; Skowron, P.; Slinko, A.; and Talmon, N. 2017. Multiwinner voting: A new challenge for social choice theory. *Trends in Computational Social Choice* 74: 27–47.

Lackner, M.; and Skowron, P. 2018. Consistent approval-based multi-winner rules. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, 47–48.

Skowron, P.; Faliszewski, P.; and Slinko, A. 2019. Axiomatic characterization of committee scoring rules. *J. Econ. Theory* 180: 244–273.

Xia, L.; and Conitzer, V. 2010. Compilation Complexity of Common Voting Rules. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, 915–920.