Strategic Voting with Incomplete Information

Ulle Endriss
Institute for Logic, Language and Computation
University of Amsterdam

joint work with Svetlana Obraztsova,
Maria Polukarov, and Jeffrey S. Rosenschein

Talk Outline

Much classical work in social choice theory assumes that strategic voters know exactly how everyone else will vote.

Instead, we assume you only have *incomplete information* and we explore the consequences of this restriction:

- effects on the *manipulability* of voting rules
- effects on the convergence of *iterative voting* processes

Preliminaries

Set of *voters* $N = \{1, \ldots, n\}$ and set of *candidates* C, with $|C| = m$.

True *preferences* $\succ_i$ and declared *ballots* b_i are linear orders in $\mathcal{L}(C)$.

Resolute *voting rule* $F : \mathcal{L}(C)^n \to C$ to pick a single winner.

To ensure resoluteness, we use *lexicographic tie-breaking*.

Focus on *Copeland* and *positional scoring rules*, including in particular *plurality*, *veto*, and other k*-approval* rules.

Safe Manipulation under Uncertainty

Information function π mapping profile $\boldsymbol{b}$ to "information" $\pi(\boldsymbol{b})$, e.g: winner information, score information, or majority graph information.

Given signal $\pi(\boldsymbol{b})$, voter i must consider these partial profiles possible:

$$\mathcal{W}_i^{\pi(\boldsymbol{b})} \quad = \quad \left\{\boldsymbol{b}'_{-i} \in \mathcal{L}(C)^{n-1} \mid \pi(b_i, \boldsymbol{b}'_{-i}) = \pi(\boldsymbol{b})\right\}$$

She might manipulate by voting $b_i^\star$ instead of b_i if both:

- $F(b_i^\star, \boldsymbol{b}^\star_{-i}) \succ_i F(b_i, \boldsymbol{b}^\star_{-i})$ *for some* $\boldsymbol{b}^\star_{-i} \in \mathcal{W}_i^{\pi(\boldsymbol{b})}$
- $F(b_i^\star, \boldsymbol{b}'_{-i}) \succcurlyeq_i F(b_i, \boldsymbol{b}'_{-i})$ *for all* $\boldsymbol{b}'_{-i} \in \mathcal{W}_i^{\pi(\boldsymbol{b})}$

Results on Manipulability

The general spirit of the *Gibbard-Satterthwaite Theorem* prevails: essentially all reasonable voting rules are susceptible to manipulation.

But we were able to identify some exceptions, such as this one:

Proposition 1 *Given majority graph information, the k-approval rules with $k \leqslant m - 2$ are immune to manipulation.*

Iterative Voting

Iterative voting with voting rule F under information function π:

- *initialise:* all voters vote truthfully $[b_i^0 := \succ_i]$
- then *repeat:* some voter i manipulates $[\boldsymbol{b}^{k+1} := (b_i^\star, \boldsymbol{b}_{-i}^k)]$

Will this process *converge*?

- to a *stable profile* (nobody wants to update anymore)?
- to a *stable outcome* (winner won't change anymore)?

Related work: for full-information case, only rules known to converge are plurality and veto (under best-response dynamics).

Convergence Results

This would not work under full information:

Theorem 2 *When voters are given only winner information, iterative Copeland voting always converges to a stable outcome.*

For positional scoring rules we need tighter assumptions:

Theorem 3 *When voters are given only winner information, iterative PSR voting always converges to a stable outcome—if voters only make minimal updates (in terms of Kendall tau distance).*

Last Slide

We have seen that restricting the information a manipulator has access to in an election can sometimes have positive effects:

- in terms of rendering a reasonable voting rule strategy-proof
- in terms of ensuring convergence of iterative voting