How should we model incomplete information in strategic voting?

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

Coping with Strategic Manipulation

According to Gibbard-Satterthwaite & Co, there are no decent voting rules that are strategyproof. *Bad!*

Strategies of the social choice theorist for coping:

- Domain restrictions
- Computational barriers
- Finding virtue in strategic manipulation after all

But maybe this is overlooking the most obvious route to salvation?

Informational Barriers to Strategic Manipulation

Unrealistic to assume the manipulator has full information regarding the others' voting intentions. *So G-S does not actually apply!*

This raises two questions:

- How should we model the information available to voters?
- What (positive) results can we get for these models?

Sampling of Preferences

Osborne and Rubinstein (2003) propose a model under which each voter knows the preferences of a *small sample* of the electorate and assumes that this sample is *representative*.

They explore this model for *plurality* voting with *three alternatives*, *single-peaked preferences*, and *samples of size 2 and 3*.

M.J. Osborne and A. Rubinstein. Sampling Equilibrium, with an Application to Strategic Voting. *Games and Economic Behavior*, 45:434–441, 2003.

Model Based on a Social Network

Chopra et al. (2004) work with a directed graph ("*knowledge graph*"): voter i knows the preferences of voter j if there is an edge from i to j.

They analyse how the structure of the graph affects convergence of *iterative voting* (where voters repeatedly manipulate).

S. Chopra, E. Pacuit, and R. Parikh. Knowledge-Theoretic Properties of Strategic Voting. Proc. JELIA-2004.

Safe Manipulation

A voter i's *friends* are the voters with the same preference order as i.

Safe manipulation $=$ manipulation that is beneficial for some number k of friends copying your move and that is not detrimental for any k.

Think of this as model of manipulation under incomplete information: each k corresponds to a possible profile.

Slinko and White (2014) show that the *Gibbard-Satterthwaite Theorem* still persists under this kind of incomplete information.

A. Slinko and S. White. Is it ever safe to vote strategically? *Social Choice and Welfare*, 43:403–427, 2014.

Set of Possible Profiles

Conitzer et al. (2011) introduce a very general model where, for every voter i, the set of *partial profiles* she deems possible is given explicitly:

$$\mathcal{W}_i \subseteq \mathcal{L}(X)^{n-1}$$

If i is *cautious*, she will *manipulate* using $\succ_i^\star$ instead of $\succ_i$ only if both:

- $F(\succ_i^\star, \boldsymbol{\succ}'_{-i}) \succ_i F(\succ_i, \boldsymbol{\succ}'_{-i})$ *for some* $\boldsymbol{\succ}'_{-i} \in \mathcal{W}_i$
- $F(\succ_i^\star, \boldsymbol{\succ}'_{-i}) \succeq_i F(\succ_i, \boldsymbol{\succ}'_{-i})$ *for all* $\boldsymbol{\succ}'_{-i} \in \mathcal{W}_i$

Conitzer et al. focus on the case where $\mathcal{W}_i$ is induced by a profile of *partial-order preferences* and study *computational barriers* to manip'n.

V. Conitzer, T. Walsh, and L. Xia. Dominating Manipulations in Voting with Partial Information. Proc. AAAI-2011.

Possible Profiles Induced by an Opinion Poll

Let π be a commonly known function, mapping any truthful profile $\boldsymbol{\succ} = (\succ_1, \ldots, \succ_n)$ to a public signal $\pi(\boldsymbol{\succ})$. Then voter i must deem possible any partial profile in this set:

$$\mathcal{W}_i^{\pi(\boldsymbol{\succ})} \quad = \quad \{\boldsymbol{\succ}'_{-i} \in \mathcal{L}(X)^{n-1} \mid \pi(\boldsymbol{\succ}) = \pi(\succ_i, \boldsymbol{\succ}'_{-i})\}$$

Example: π might be an *opinion poll* that returns, say, the winner of the election, or the plurality score of every alternative.

If i is cautious, she will manipulate using $\succ_i^\star$ instead of $\succ_i$ only if both:

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

A. Reijngoud and U. Endriss. Voter Response to Iterated Poll Information. Proc. AAMAS-2012.

Example: Manipulation under Zero Information

When does lack of information constitute a barrier to manipulation?

Contrary to intuition, some (strange) voting rules *can be manipulated* even if you have *no information* at all:

> Take the voting rule that elects the Condorcet winner if it exists, and otherwise the bottom alternative of voter 1.

If voter 1's true preferences are $x \succ_1 y \succ_1 z$, she can never do worse by voting $x \succ z \succ y$, and she does better if the others vote like this:

$$x \succ_2 y \succ_2 z$$

$$y \succ_3 x \succ_3 z$$

$$y \succ_4 x \succ_4 z$$

But it really is a strange voting rule . . .

Open: *Which rules are immune to manipulation under zero information?*

Antiplurality and Winner Information

Let n be the number of voters and let m be the number of alternatives.

Theorem 1 (Reijngoud and Endriss, 2012) *For $n \geqslant 2m - 2$, the antiplurality rule (with ties getting broken lexicographically) is strategyproof if voters only know the winner for the truthful profile.*

Proof: Suppose $m \geqslant 3$ (other cases: clear). Consider voter i. Let x be voter i's *worst alternative*. Let $x^\star$ be the truthful *winner*.

Distinguish two cases:

- $x = x^\star$: Nothing she can do to change the outcome. ✓
- $x \neq x^\star$: If i manipulates by vetoing some $y \neq x$ (possibly $y = x^\star$), then x gains a point and $x^\star$ does not, so x *could* now win. ✓

A. Reijngoud and U. Endriss. Voter Response to Iterated Poll Information. Proc. AAMAS-2012.

K-Approval and Majority-Graph Information

One of the very few other positive results of this kind:

Theorem 2 (Endriss et al., 2016) *For a sufficiently large number of voters, all k-approval rules (except for antiplurality) are strategyproof if voters only know the majority graph for the truthful profile.*

Proof: Bit complicated. ✓

U. Endriss, M. Polukarov, S. Obraztsova, and J.S. Rosenschein. Strategic Voting with Incomplete Information. Proc. IJCAI-2016.

Distance-Based Uncertainty

Another option is to assume that, for some notion of distance d, each voter i deems possible the set of partial profiles in the neighbourhood of some specific such profile $\succ_{-i}$:

$$\mathcal{W}_i^{d,k,\succ_{-i}} \;=\; \{\succ'_{-i} \in \mathcal{L}(X)^{n-1} \mid d(\succ_{-i}, \succ'_{-i}) \leqslant k\}$$

Meir et al. (2014) use this for the special case of the *plurality rule* (and "anonymous" profiles, i.e., lists of scores).

R. Meir, O. Lev, and J.S. Rosenschein. A Local-Dominance Theory of Voting Equilibria. Proc. EC-2014.

Borda and Swap-Distance Uncertainty

Now use the *swap distance* to measure distances between profiles.
Let n be the number of voters and let m be the number of alternatives.

Theorem 3 (Damanik et al., 2017) *The Borda rule is strategyproof if voters only know that the truthful profile belongs to a swap-distance neighbourhood with radius $2nm$.*

Proof: Suppose voter i contemplates untruthfully ranking y above x.
But this partial profile is possible (within $2nm$ swaps):

half of others: $x \succ y \succ \cdots$
half of others: $y \succ x \succ \cdots$

So voter i risks getting y elected instead of x. ✓

R. Damanik, S. Ionescu, L. van Oostveen, and P. Seip. Strategic Manipulation for Voting under Distance-Based Uncertainty. University of Amsterdam, 2017.

Last Slide

I have reviewed a number of proposals for modelling the incomplete information available to a strategic voter during an election:

- sampling (Osborne and Rubinstein, 2003)
- knowledge graph (Chopra, Pacuit, and Parikh, 2004)
- safe manipulation (Slinko and White, 2014)
- set of possible profiles (Conitzer, Walsh, and Xia, 2011)
- possible profiles induced by public signal (Reijngoud and E., 2012)
- distance-based uncertainty (Meir, Lev, and Rosenschein, 2014)

What is the right model for this problem?

And how does manipulation work exactly under incomplete information? (I have only discussed the "cautious" approach.)