Computational Social Choice 2021

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

[http://www.illc.uva.nl/~ulle/teaching/comsoc/2021/]

Plan for Today

It is not always in the best interest of voters to truthfully reveal their preferences when voting. This is called *strategic manipulation*.

We'll prove a seminal result, the *Gibbard-Satterthwaite Theorem*, that shows that can't be avoided: (essentially) strategyproof $\Rightarrow$ dictatorial

We then will review three approaches for addressing the challenges raised by strategic manipulation:

- *Domain restrictions:* excluding problematic profiles
- *Computational barriers:* making manipulation intractable
- *Informational barriers:* hiding information from manipulators

Example

Recall that under the *plurality rule* (used in most political elections) the candidate ranked first most often wins the election.

Assume the preferences of the people in, say, Florida are as follows:

49%:	Bush $\succ$ Gore $\succ$ Nader
20%:	Gore $\succ$ Nader $\succ$ Bush
20%:	Gore $\succ$ Bush $\succ$ Nader
11%:	Nader $\succ$ Gore $\succ$ Bush

So even if nobody is cheating, Bush will win this election.

It would have been in the interest of the Nader supporters to manipulate, i.e., to misrepresent their preferences.

Is there a better voting rule that avoids this problem?

Truthfulness, Manipulation, Strategyproofness

Today, we only deal with *resolute* voting rules $F : \mathcal{L}(A)^n \to A$.

Unlike for all earlier results discussed, we now have to distinguish:

- the *ballot* a voter reports
- her actual *preference* order

Both are elements of $\mathcal{L}(A)$. If they coincide, then the voter is *truthful*.

F is *strategyproof* (or *immune to manipulation*) if for no voter $i \in N$ there exist a profile $\boldsymbol{R}$ (including i's *truthful preference* R_i) and an *untruthful ballot* R'_i for i such that R_i ranks $F(R'_i, \boldsymbol{R}_{-i})$ above $F(\boldsymbol{R})$.

Thus: Nobody has an incentive to misrepresent their preferences.

Notation: $(R'_i, \boldsymbol{R}_{-i})$ is the profile obtained by replacing R_i in $\boldsymbol{R}$ by R'_i.

Importance of Strategyproofness

Why do we want voting rules to be strategyproof?

- "Thou shalt not bear false witness against thy neighbour."
- Voters should not have to waste resources pondering over what other voters will do and trying to figure out how best to respond.
- If everyone strategises (and makes mistakes when guessing how others will vote), then the final ballot profile will be very far from the electorate's true preferences and thus the election winner may not be representative of their wishes at all.

The Gibbard-Satterthwaite Theorem

Recall: A resolute SCF F is *surjective* if for every alternative $x \in A$ there exists a profile $\boldsymbol{R}$ such that $F(\boldsymbol{R}) = x$.

Gibbard (1973) and Satterthwaite (1975) independently proved:

Theorem 1 (Gibbard-Satterthwaite) *Any resolute SCF for $\geqslant 3$ alternatives that is surjective and strategyproof is a dictatorship.*

Remarks:

- a *surprising* result + not applicable in case of *two* alternatives
- The opposite direction is clear: *dictatorial* $\Rightarrow$ *strategyproof*
- *Random* rules don't count (but might be 'strategyproof').

A. Gibbard. Manipulation of Voting Schemes: A General Result. *Econometrica*, 1973.

M.A. Satterthwaite. Strategy-proofness and Arrow's Conditions. *Journal of Economic Theory*, 1975.

Proof

We shall prove the Gibbard-Satterthwaite Theorem to be a corollary of the Muller-Satterthwaite Theorem (even if, historically, G-S came first).

Recall the *Muller-Satterthwaite Theorem:*

- Any *resolute* SCF for $\geqslant 3$ alternatives that is *surjective* and *strongly monotonic* must be a *dictatorship*.

We shall prove a lemma showing that strategyproofness implies strong monotonicity (and we'll be done). ✓ (Details are in my review paper.)

For other short proofs of G-S, see Barberà (1983) and Benoît (2000).

S. Barberà. Strategy-Proofness and Pivotal Voters: A Direct Proof the Gibbard-Satterthwaite Theorem. *International Economic Review*, 1983.

J.-P. Benoît. The Gibbard-Satterthwaite Theorem: A Simple Proof. *Economic Letters*, 2000.

U. Endriss. Logic and Social Choice Theory. In A. Gupta and J. van Benthem (eds.), *Logic and Philosophy Today*, College Publications, 2011.

Strategyproofness implies Strong Monotonicity

Lemma 2 *Any resolute SCF that is strategyproof (SP) must also be strongly monotonic (SM).*

- SP: no incentive to vote untruthfully
- SM: $F(\boldsymbol{R}) = x \;\Rightarrow\; F(\boldsymbol{R'}) = x$ if $N^{\boldsymbol{R}}_{x \succ y} \subseteq N^{\boldsymbol{R'}}_{x \succ y}$ for all y

Proof: We'll prove the contrapositive. So assume F is *not* SM.
So there exist $x, x' \in A$ with $x \neq x'$ and profiles $\boldsymbol{R}, \boldsymbol{R'}$ such that:

- $N^{\boldsymbol{R}}_{x \succ y} \subseteq N^{\boldsymbol{R'}}_{x \succ y}$ for all alternatives y, including x' $(\star)$
- $F(\boldsymbol{R}) = x$ and $F(\boldsymbol{R'}) = x'$

Moving from $\boldsymbol{R}$ to $\boldsymbol{R'}$, there must be a *first* voter affecting the winner.
So w.l.o.g., assume $\boldsymbol{R}$ and $\boldsymbol{R'}$ differ only w.r.t. voter i. Two cases:

- $i \in N^{\boldsymbol{R'}}_{x \succ x'}$: if i's true preferences are as in $\boldsymbol{R'}$, she can benefit from voting instead as in $\boldsymbol{R}$ $\Rightarrow$ F is not SP ✓
- $i \notin N^{\boldsymbol{R'}}_{x \succ x'} \;\Rightarrow^{(\star)}\; i \notin N^{\boldsymbol{R}}_{x \succ x'} \;\Rightarrow\; i \in N^{\boldsymbol{R}}_{x' \succ x}$: if i's true preferences are as in $\boldsymbol{R}$, she can benefit from voting as in $\boldsymbol{R'}$ $\Rightarrow$ F is not SP ✓

Resoluteness Assumption

Note that the Gibbard-Satterthwaite Theorem applies to *resolute* rules, while almost all voting rules we have a name for actually are *irresolute*.

Does this provide a way out? Not really:

- For most applications we really need a *single winner* in the end, so our definition of F must incorporate a *tie-breaking rule*.
- There are (only slightly more favourable) impossibility theorems for irresolute rules as well, notably the *Duggan-Schwartz Theorem*.

J. Duggan and T. Schwartz. Strategic Manipulation w/o Resoluteness or Shared Beliefs: Gibbard-Satterthwaite Generalized. *Social Choice and Welfare*, 2000.

The Bigger Picture

We have by now seen three impossibility theorems for *resolute* SCF's, all of which apply in case there are at least *three alternatives:*

Gibbard-Satterthwaite Theorem
[surjective + strategyproof $\Rightarrow$ dictatorial]

$\Uparrow$

Muller-Satterthwaite Theorem
[surjective + strongly monotonic $\Rightarrow$ dictatorial

$\Uparrow$

Arrow's Theorem
[Paretian + independent $\Rightarrow$ dictatorial]

We proved Arrow's Theorem by analysing when a coalition can force a pairwise ranking. The other two results followed by comparing axioms.

Barriers to Strategic Manipulation

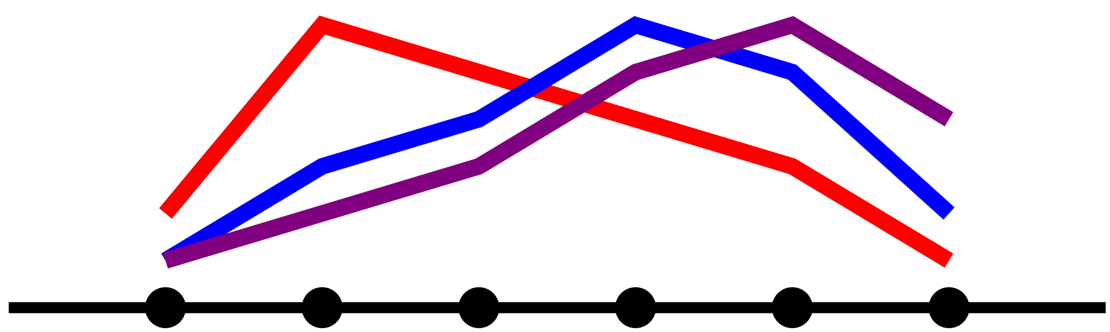
Is that the end of it? No! Next we are going to briefly review three kinds of barriers against strategic manipulation . . .

Domain Restriction: Single-Peaked Preferences

Every voting rule can be manipulated, but not in all profiles. Can we do better if we restrict attention to specific (natural) profiles?

We only discuss the oldest and most famous domain restriction . . .

A profile $\boldsymbol{R} = (R_1, \ldots, R_n)$ is *single-peaked* if we can arrange the alternatives from left to right along some dimension $\gg$ such that R_i ranks x above y whenever x is between y and $\mathrm{top}(R_i)$ according to $\gg$.



Sometimes a natural assumption: traditional political parties, agreeing on a number (e.g., legal drinking age), . . .

Strategyproofness of the Median-Voter Rule

For a given dimension $\gg$, the *median-voter rule* asks each voter for her top alternative and elects the alternative proposed by the voter corresponding to the median w.r.t. $\gg$.

Theorem 3 *If an odd number of voters have preferences that are single-peaked w.r.t.* $\gg$*, then the median-voter rule is strategyproof.*

Proof: W.l.o.g., our manipulator's top alternative is *to the right* of the median (the winner). If she declares a peak further to the right, nothing will change. If she declares a peak further to the left, either nothing will change, or the new winner will be even worse. ✓

This is closely related to Black's *Median Voter Theorem*, showing that under the same conditions a Condorcet winner exists and is elected.

D. Black. On the Rationale of Group Decision-Making. *The Journal of Political Economy*, 1948.

Computational Barriers to Manipulation

Every voting rule can be manipulated in some profiles. But even when it is *possible* to manipulate, maybe actually doing so is *difficult*?

Tools from *complexity theory* can help make this idea precise.

> *If manipulation is computationally intractable for F, then we might consider F resistant (but not immune) to manipulation.*

Does not work for most rules, but STV manipulation is NP-hard.

Discussion: Practical significance of these results is debatable, in particular when they presuppose that there are many alternatives.

J.J. Bartholdi III, C.A. Tovey, and M.A. Trick. The Computational Difficulty of Manipulating an Election. *Social Choice and Welfare*, 1989.

V. Conitzer and T. Walsh. Barriers to Manipulation in Voting. In F. Brandt et al. (eds.), *Handbook of Computational Social Choice*. CUP, 2016.

Coalitional Manipulation

It rarely is the case that a *single* voter really can make a difference. So we should look into *manipulation by a coalition* of voters.

Variants of the problem:

- Ballots may be *weighted* or *unweighted*.

  Examples: countries in the EU, shareholders of a company

- Manipulation may be *constructive* (making alternative x win) or *destructive* (ensuring x does not win).

Decision Problems

We consider two decision problems, for a given voting rule F:

ConstructiveManipulability(F)
Input: List of weighted ballots; set of weighted manipulators; $x \in A$.
Question: Are there ballots for the manipulators such that x wins?

DestructiveManipulability(F)
Input: List of weighted ballots; set of weighted manipulators; $x \in A$.
Question: Are there ballots for the manipulators such that x loses?

Constructive Manipulation under Borda

In the context of coalitional manipulation with weighted voters, we can get hardness results for elections with small numbers of alternatives:

Theorem 4 (Conitzer et al., 2007) *For the Borda rule, the constructive coalitional manipulation problem with weighted voters is NP-complete for $\geqslant 3$ alternatives.*

Proof: We have to prove NP-membership and NP-hardness:

- NP-membership: easy (if you guess ballots for the manipulators, we can check that it works in polynomial time)
- NP-hardness: for three alternatives by reduction from PARTITION (next slide); hardness for more alternatives follows

V. Conitzer, T. Sandholm, and J. Lang. When are Elections with Few Candidates Hard to Manipulate? *Journal of the ACM*, 54(3), Article 14, 2007.

Proof of NP-hardness

We use a reduction from the NP-complete PARTITION problem:

PARTITION
Input: $(w_1, \ldots, w_n) \in \mathbb{N}^n$
Question: Is there a set $S \subseteq \{1, \ldots, n\}$ s.t. $\sum_{i \in S} w_i = \frac{1}{2} \sum_{i=1}^{n} w_i$?

Let $K := \sum_{i=1}^{n} w_i$. Given an instance of PARTITION, we construct an election with $n + 2$ weighted voters and three alternatives:

- two voters with weight $\frac{1}{2}K - \frac{1}{4}$, voting $(a \succ b \succ c)$ and $(b \succ a \succ c)$
- a coalition of n voters with weights $w_1, \ldots, w_n$ who want c to win

Clearly, each manipulator should vote either $(c \succ a \succ b)$ or $(c \succ b \succ a)$. Suppose there does exist a partition. Then they can vote like this:

- manipulators corresponding to elements in S vote $(c \succ a \succ b)$
- manipulators corresponding to elements outside S vote $(c \succ b \succ a)$

Scores: $2K$ for c; $\frac{1}{2}K + (\frac{1}{2}K - \frac{1}{4}) \cdot (2 + 1) = 2K - \frac{3}{4}$ for both a and b
If there is no partition, then either a or b will get at least 1 point more.
Hence, manipulation is feasible <u>iff</u> there exists a partition. ✓

Destructive Manipulation under Borda

Theorem 5 (Conitzer et al., 2007) *For the Borda rule, the destructive coalitional manip. problem with weighted voters is in P.*

Proof: Let x be the alternative the manipulators want to lose.

For every $y \neq x$, simply try everyone ranking y at the top and x at the bottom. If none of these $m-1$ attempts work, nothing will. ✓

V. Conitzer, T. Sandholm, and J. Lang. When are Elections with Few Candidates Hard to Manipulate? *Journal of the ACM*, 54(3), Article 14, 2007.

Informational Barriers to Manipulation

Suppose voter i has only *partial information* about the profile. If π is a function mapping any truthful profile $\boldsymbol{R}$ to the information $\pi(\boldsymbol{R})$ given to i, then i must consider possible any profile in this set:

$$\mathcal{W}_i^{\pi(\boldsymbol{R})} \quad = \quad \{\, \boldsymbol{R}' \in \mathcal{L}(A)^n \mid \pi(\boldsymbol{R}) = \pi(\boldsymbol{R}') \text{ and } R_i = R'_i \,\}$$

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

Now i will manipulate using R'_i only if doing so is *strictly better* for her in at least one profile in $\mathcal{W}_i^{\pi(\boldsymbol{R})}$ and *not worse* in any of the others.

Limited positive results to date. For instance, the *antiplurality rule* is strategyproof when voters only have *winner information*.

Remark: Interesting, still very much underexplored research direction.

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

Summary

We saw that *strategic manipulation* is a major problem in voting:

- *Gibbard-Satterthwaite:* SP + surjectivity $\Rightarrow$ dictatorship

But we also saw that there are approaches for tackling this problem:

- *Domain restrictions*
- *Computational barriers*
- *Informational barriers*

The study of strategic manipulation is very much at the intersection of social choice theory with *game theory* and *mechanism design*.

Other forms of strategic behaviour that may occur in the context of elections include *bribery* and *gerrymandering*.

What next? Moving away from the classical model of voting, we will start looking into new ideas for democratic decision making.