Computational Social Choice 2026

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

http://www.illc.uva.nl/~ulle/teaching/comsoc/2026/

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*, showing this can't be avoided: (essentially) *strategyproof* $\Rightarrow$ *dictatorial*

We'll also see how to generalise this result from resolute to *irresolute* voting rules, and we'll briefly discuss other forms of *strategic behaviour*.

The Problem of Strategic Manipulation

One requirement we might have is that we don't want voters to have an incentive to misrepresent their true preferences.

Remember what happened in Florida in 2000 (*stylised*):

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

Under *plurality*, Bush will win. Nader supporters had an incentive to pretend they prefer Gore. We say: *Plurality is not strategyproof*.

Exercise: *Is there a better voting rule that avoids this problem?*

Truthfulness, Manipulation, Strategyproofness

For now, we focus on *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 doing so), the ballot profile will be far from the electorate's true preferences and thus the election winner may not be representative of their wishes at all.

The Full-Information Assumption

Here, as in most work on the topic, we make the assumption that the manipulator has *full information* about the ballots of the other voters.

Is this always realistic? No. But:

- In *small committees* (e.g., members of a department voting on who to hire) the full-information assumption is fairly realistic.
- Even in large political elections *poll information* may be accurate enough to allow groups of voters (though not individuals) to perform similar acts of manipulation as discussed here.
- When looking for *protection against manipulation*, we should assume the *worst case*, where the manipulator has full information.

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:

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

Remark: *Random* rules don't count (but might be 'strategyproof').

Exercise: *Show that the theorem does not hold for $m = 2$.*

Exercise: *Prove the dual: dictatorial $\Rightarrow$ surjective + 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 $m \geqslant 3$ 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). Full details 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

Claim: strategyproof (SP) $\Rightarrow$ 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' $\quad(\star)$
- $F(\boldsymbol{R}) = x$ 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 ✓

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.

Shortcomings of Resolute Voting Rules

The Gibbard-Satterthwaite Theorem only applies to *resolute* rules.
But the restriction to resolute rules is problematic:

- No 'natural' voting rule is resolute (w/o tie-breaking rule).
- We can get very basic impossibilities for resolute rules:

  We've seen already that *no resolute* voting rule for *two voters* and *two alternatives* can be both *anonymous* and *neutral*.

So maybe we should be analysing *irresolute* voting rules instead?

Manipulability w.r.t. Psychological Assumptions

To analyse manipulability when we might get a set of winners, we need to make assumptions on how voters rank *sets of alternatives*, e.g.:

- A voter is an *optimist* if she prefers X over Y whenever she prefers her favourite $x \in X$ over her favourite $y \in Y$.
- A voter is a *pessimist* if she prefers X over Y whenever she prefers her least preferred $x \in X$ over her least preferred $y \in Y$.

Now we can speak about manipulability by certain types of voters:

- F is called *immune to manipulation by optimistic voters* if no optimistic voter can ever benefit from voting untruthfully.
- F is called *immune to manipulation by pessimistic voters* if no pessimistic voter can ever benefit from voting untruthfully.

Remark: All of this is an application of the '*ranking sets of objects*' framework briefly discussed in the first lecture.

Axiom: Nonimposition

Let F be an *irresolute* SCF $F : \mathcal{L}(A)^n \to 2^A \setminus \{\emptyset\}$.

- F is *nonimposed* if for every alternative x there exists a profile $\boldsymbol{R}$ under which x is the unique winner: $F(\boldsymbol{R}) = \{x\}$.

For comparison, *surjectivity* means that for every element in the co-domain of F there is an input producing that element. Thus:

$$\text{resolute} \Rightarrow (\text{nonimposed} = \text{surjective})$$

Dictatorships for Irresolute Rules

Let F be an *irresolute* SCF $F : \mathcal{L}(A)^n \to 2^A \setminus \{\emptyset\}$.

There are two natural notions of dictatorship for such rules:

- Voter $i \in N$ is called a (strong) *dictator* if $F(\boldsymbol{R}) = \{\mathrm{top}(R_i)\}$ for every profile $\boldsymbol{R} \in \mathcal{L}(A)^n$.
- Voter $i \in N$ is called a *weak dictator* if $\mathrm{top}(R_i) \in F(\boldsymbol{R})$ for every profile $\boldsymbol{R} \in \mathcal{L}(A)^n$. (Such a voter is also called a *nominator*.)

F is called *weakly dictatorial* if it has (at least) a weak dictator.
Otherwise F is called *strongly nondictatorial*.

The Duggan-Schwartz Theorem

There are several extensions of the Gibbard-Satterthwaite Theorem to irresolute voting rules. The Duggan-Schwartz Theorem is widely regarded as the most important of these results.

Our statement of the theorem follows Taylor (2002):

Duggan-Schwartz Theorem: *Any voting rule for $m \geqslant 3$ alternatives that is nonimposed and immune to manipulation by both optimistic and pessimistic voters must be weakly dictatorial.*

Proof: Omitted.

The Gibbard-Satterthwaite Theorem is a corollary. Exercise: *Why?*

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

A.D. Taylor. The Manipulability of Voting Systems. *The American Mathematical Monthly*, 2002.

Other Forms of Strategic Behaviour

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

Summary

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

- *Gibbard-Satterthwaite:* SP + surjectivity $\Rightarrow$ dictatorship
- *Duggan-Schwartz:* dropping resoluteness does not help much

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

What next? Different ways of coping with strategic manipulation.