Topics in Computational Social Choice 2026

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

Plan for Today

- Discussion of second exercise sheet
- Strategic manipulation and the Gibbard-Satterthwaite Theorem
- Brief mentioning of various research methods used in COMSOC
- Group exercise on identifying research questions

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

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.

Axiom: Surjectivity

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

So no alternative is barred from winning at the outset.

The Gibbard-Satterthwaite Theorem

After many years of people trying to find a voting rule that cannot be manipulated, Gibbard (1973) and Satterthwaite (1975) 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: *Explain why surjectivity is needed as a condition.*

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 Plan

We will prove the *G-S Theorem* as a corollary to *Arrow's Theorem*.

We are done if we can show that *surjectivity* and *strategyproofness* together imply both *Pareto efficiency* and *independence*. (*Why?*)

Deriving Independence (1)

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' $(\star)$
- $F(\boldsymbol{R}) = x$ 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 ✓

Deriving Independence (2)

Recall: F is *independent* if, for $x \neq y$, we have that $N^{\boldsymbol{R}}_{x \succ y} = N^{\boldsymbol{R}'}_{x \succ y}$ and $F(\boldsymbol{R}) = x$ together imply $F(\boldsymbol{R}') \neq y$.

Claim: strongly monotonic (SM) $\Rightarrow$ independent

Proof: Suppose F is SM, $x \neq y$, $N^{\boldsymbol{R}}_{x \succ y} = N^{\boldsymbol{R}'}_{x \succ y}$, and $F(\boldsymbol{R}) = x$.

Construct a third profile $\boldsymbol{R}''$:

- All individuals rank x and y in the top-two positions.
- The relative rankings of x vs. y are as in $\boldsymbol{R}$, i.e., $N^{\boldsymbol{R}''}_{x \succ y} = N^{\boldsymbol{R}}_{x \succ y}$.
- Rest: whatever

By strong monotonicity, $F(\boldsymbol{R}) = x$ implies $F(\boldsymbol{R}'') = x$.
By strong monotonicity, $F(\boldsymbol{R}') = y$ *would* imply $F(\boldsymbol{R}'') = y$.
Thus, we must have $F(\boldsymbol{R}') \neq y$. ✓

Deriving the Pareto Principle

Recall: F is *Paretian* if $N^{\boldsymbol{R}}_{x \succ y} = N$ implies $F(\boldsymbol{R}) \neq y$.

Claim: If F is surjective and strategyproof, then F is also Paretian.

Proof: Suppose F is surjective and SP (so also independent and SM).

Take any two alternatives x and y.

From surjectivity: x will win for *some* profile $\boldsymbol{R}$.

Starting in $\boldsymbol{R}$, have everyone move x above y (if not above already).
From SM: x still wins.

From independence: y does not win for *any* profile where all individuals rank $x \succ y$. ✓

Reflection on Impossibility Results

At this point, we have seen impossibilities triggered by these axioms:

- Anonymity + Neutrality (for resolute rules only)
- Arrow: Pareto + Independence + Nondictatorship
- G-S: Surjectivity + Strategyproofness + Nondictatorship

Exercise: *How should we interpret these technical results?*

Barriers to Strategic Manipulation

By the G-S Theorem, designing a (not obviously terrible) voting rule that is immune to strategic manipulation is impossible.

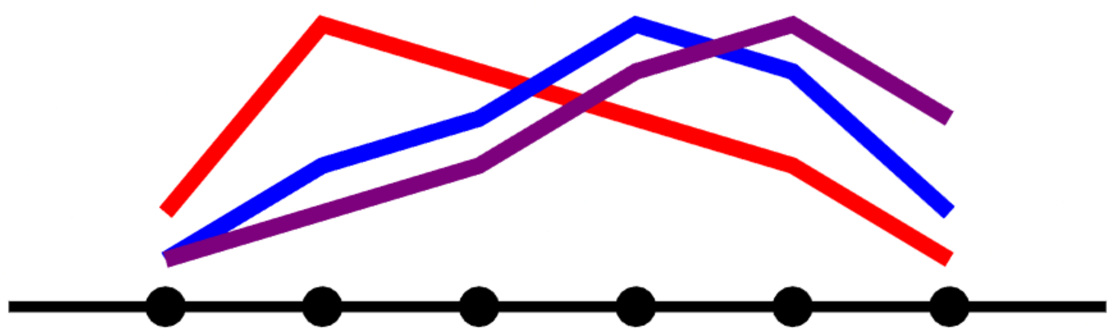
Is that the end of it? No! Next we are going to briefly review three kinds of barriers against strategic manipulation . . .

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, spatial voting, 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. $\gg$.

Folklore Theorem: *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 profile. 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 IRV/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.

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 *no worse* in any of the others.

Limited positive results to date. One is the insight that the *veto 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.

Other Topics in Voting

Let us briefly review some other research topics in COMSOC related to voting, highlighting the methodological variety of the field.

Bribery and Control

Manipulation is about strategic conduct by the voters.

Strategic conduct by the election chair or outside parties:

- *Bribery:* Can we swing the election by bribing up to k voters?
- *Control:* Can we swing the election by adding/deleting voters?

Most work concerns questions of *computational complexity*.

P. Faliszewski and J. Rothe. Control and Bribery in Voting. In: F. Brandt et al. (eds.), *Handbook of COMSOC*, Cambridge University Press, 2016.

The Condorcet Jury Theorem

May's Theorem offers good reasons for using the simple majority rule.

But the simple majority rule is also *epistemically* attractive, in terms of *tracking the truth* (assuming there is a "correct" choice):

Condorcet-Jury Theorem (1795): *Suppose a jury of n voters need to select the better of two alternatives and each voter independently makes the correct decision with the same probability $p > \frac{1}{2}$. Then the probability that the simple majority rule returns the correct decision increases monotonically in n and approaches 1 as n goes to infinity.*

Proof sketch: By the law of large numbers, the number of voters making the correct choice approaches $p \cdot n > \frac{1}{2} \cdot n$. ✓

Writings of the Marquis de Condorcet. In I. McLean and A. Urken (eds.), *Classics of Social Choice*. University of Michigan Press, 1995.

Iterative Voting

Imagine voters can update their ballot after observing the other ballots.

Example: Doodle polls

Suppose voters can update again and again, one at a time.

- What are reasonable behavioural assumptions?
- Will this converge?
- If it does, what can we say about the final profile?

For a natural game-theoretical model regarding voter behaviour, we get convergence for *plurality* and *veto*, but no other standard voting rule.

R. Meir. Iterative Voting. In U. Endriss (ed.), *Trends in Computational Social Choice*. AI Access, 2017.

Possible Election Winners

Suppose we currently only have partial information about the voter preferences (but they themselves do have complete preferences).

We may ask: what are the *possible winners* under voting rule F?

Why is this interesting?

- Postal ballots may arrive late
- Some alternatives available only after voting has started
- Elicitation: can stop once *possible winner* = *necessary winner*
- Relationship to (coalitional) manipulation (Exercise: *What is it?*)

Lots of complexity-theoretic results in the literature.

Example: General case in P for *plurality* but NP-complete for *Borda*.

K. Konczak and J. Lang. Voting Procedures with Incomplete Preferences. Proc. Multidisciplinary Workshop on Advances in Preference Handling 2005.

Compilation Complexity

For voting rule F, given a partial profile, how many bits do we need to store (in the worst case) so we can compute the outcome later on, once the profile is complete? This is the *compilation complexity* of F.

Let n be the number of (old) voters and m the number of alternatives. We need $\lceil \log m \rceil$ bits to represent the name of one alternative. So the CC of any rule F is at most $n\lceil \log(m!) \rceil$ (just store all ballots!).

Exercise: *Prove these following observations to be correct!*

- for *anonymous* rule is at most $\min\{n\lceil \log(m!) \rceil, m!\lceil \log(n+1) \rceil\}$
- for *dictatorial* rules it is $\lceil \log m \rceil$
- for *constant* rules (always electing the same winner) it is 0

Chevaleyre et al. (2009) establish further such results, e.g., the fact that the CC of *Borda* is $\Theta(m \log(nm))$. Exercise: *Upper bound clear?*

Y. Chevaleyre, J. Lang, N. Maudet, and G. Ravailly-Abadie. Compiling the Votes of a Subelectorate. IJCAI-2009.

Liquid Democracy and Social Networks

Liquid Democracy is the idea of allowing voters to choose between voting and delegating their votes to others (in a transitive manner).

Used in practice (famous example: *Piratenpartei* in Germany), but development of sound theoretical foundations still lacking.

The entire methodological battery of COMSOC can be brought to bear on the study of such novel models of collective decision making.

<u>More Generally:</u> COMSOC + Social Network Theory is promising.

M. Brill. Interactive Democracy. AAMAS-2018 Blue Sky Ideas Track.

U. Grandi. Social Choice and Social Networks. In U. Endriss (ed.), *Trends in Computational Social Choice*. AI Access, 2017.

Logic for Social Choice Theory

It can be insightful to model SCT problems in logic (Pauly, 2008):

- One research direction is to explore how far we can get using a *standard logic*, such as classical FOL. Do we need second-order constructs to capture IIA? (Grandi and Endriss, 2013)
- Another direction is to design *tailor-made logics* specifically for SCT (for instance, a modal logic). Can we cast the proof of Arrow's Theorem in natural deduction? (Ciná and Endriss, 2016)

M. Pauly. On the Role of Language in Social Choice Theory. *Synthese*, 2008.

U. Grandi and U. Endriss. First-Order Logic Formalisation of Impossibility Theorems in Preference Aggregation. *Journal of Philosophical Logic*, 2013.

G. Ciná and U. Endriss. Proving Classical Theorems of Social Choice Theory in Modal Logic. *Journal of Autonomous Agents and Multiagent Systems*, 2016.

Automated Reasoning for Social Choice Theory

Some work on *model checking* implementations of voting rules and verifying proofs of known results via *interactive proof assistants*.

Maybe most exiting is the use of *SAT solvers* to prove theorems in SCT. To prove an *impossibility theorem*, proceed as follows:

- model your social choice scenario using (some kind of) logic
- for fixed small values of n and m, rewrite in propositional logic
- use a SAT solver to establish unsatisfiability for these small values
- prove that impossibilities remain when we increase n or m

Results include: new proofs of known results, sharpening of known results (relaxation of conditions on n or m), and entirely new results. For simple domains, even automated *discovery* of results is feasible.

Limitation: So far mostly restricted to impossibility theorems.

C. Geist and D. Peters. Computer-Aided Methods for Social Choice Theory. In U. Endriss (ed.), *Trends in Computational Social Choice*. AI Access, 2017.

Identifying a Research Question

Option 1: Start from a real-world phenomenon that interests you.

- Come up with a model that is simple enough to permit analysis.
- Explore which of the techniques discussed might apply.

Option 2: Start from a paper you like.

- Theoretical paper? $\longrightarrow$ Design corresponding experiment!
- Experimental paper? $\longrightarrow$ Prove edge cases analytically!
- Econ paper? $\longrightarrow$ Look for algorithmic questions!
- Purely computational paper? $\longrightarrow$ Look for normative foundations!
- Negative results? $\longrightarrow$ Do better by focusing on special cases!
- General: What happens if you change some of the assumptions?

Group Assignment 1

We so far have modelled ballots as strict rankings of the alternatives. Propose an *alternative ballot format* and argue for its usefulness.

Group Assignment 2

For our standard model of voting rules aggregating lists of rankings into nonempty sets, think of a possible *application outside of politics*.

Formulate a novel *axiom* that's specifically relevant to your application.

Group Assignment 3

Give an example for a *technique* you learned about in another course that could potentially be applied in computational social choice.