Computational Social Choice: Voting Theory, Automated Reasoning, Explainability

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

What is Computational Social Choice?

Social choice theory is about methods for *collective decision making*, such as *political* decision making by groups of *economic* agents.

Its methodology ranges from the *philosophical* to the *mathematical*. It is traditionally studied in *Economics* and *Political Science* and it is a close cousin of both *decision theory* and *game theory*.

Its findings are relevant to multiple *applications*, such as these:

- How to fairly allocate resources to the members of a society?
- How to fairly divide computing time between several users?
- How to elect a president given people's preferences?
- How to combine the website rankings of multiple search engines?
- How to aggregate the views of different judges in a court case?
- How to extract information from noisy crowdsourced data?

Computational social choice, the topic of this tutorial, emphasises the fact that any method of decision making is ultimately an *algorithm*.

Why is this AI?

Ideas from Economics entered AI when it became clear that we can use them to study interaction between agents in a *multiagent system*. Nowadays, the study of so-called *economic paradigms* is all over AI.

The influential *One Hundred Year Study on Artificial Intelligence* (2016) singles out the following eleven "*hot topics*" in AI:

large-scale machine learning | deep learning | reinforcement learning
robotics | computer vision | natural language processing
collaborative systems | crowdsourcing and human computation
algorithmic game theory and *computational social choice*
internet of things | neuromorphic computing

And indeed, while COMSOC transcends several disciplines, about half of it gets published in AI conference proceedings and journals.

P. Stone et al. "Artificial Intelligence and Life in 2030". One Hundred Year Study on Artificial Intelligence. Stanford, 2016.

Plan for this Tutorial

We shall focus on the most basic scenario of *voting:* each *agent* ranks the available *alternatives* and we need to pick one of them.

Today (classical foundations):

- examples for *voting rules*
- introduction to the *axiomatic method*

Tomorrow (specific research directions):

- *automated reasoning* for social choice
- *explainability* in social choice

Three Voting Rules

Suppose n *agents* (a.k.a. *voters*) choose from a set of m *alternatives* by stating their preferences as *linear orders* over the alternatives.

How do we decide which alternative to select as the collective choice?

Here are three *voting rules* (there are many more) we might use:

- *Plurality:* elect the alternative ranked first most often
- *Plurality with runoff:* run a plurality election and retain the two front-runners; then run a majority contest between them
- *Borda:* award $m-k$ points to an alternative for every agent who ranks it in the kth position (highest score wins)

Exercise: *Do you know real-world elections where these rules are used?*

Example: Choosing a Beverage for Lunch

Consider this scenario, with nine *agents* having to choose from three *alternatives* (namely what beverage to order for a common lunch):

2 Germans:	Beer	$\succ$	Wine	$\succ$	Milk
3 French people:	Wine	$\succ$	Beer	$\succ$	Milk
4 Dutch people:	Milk	$\succ$	Beer	$\succ$	Wine

Recall that we saw three different voting rules:

- Plurality
- Plurality with runoff
- Borda

Exercise: *For each of the rules, which beverage wins the election?*

Picking a Voting Rule

So: Lots of rules. *How do you pick one?* Criteria we might use:

- *normative* requirements
- *epistemic* requirements
- *computational* requirements
- *informational* requirements

Today, we shall mainly focus on the first family of requirements.

The Axiomatic Method

The classical approach to choosing what voting rule to use, developed largely in Economics, is the so-called *axiomatic method*:

- identify normatively appealing properties of rules
- cast those properties into mathematically rigorous definitions
- explore the consequences of the thus defined "axioms"

The definitions of axioms on the following slides are only sketched, but can be made mathematically precise (see paper cited below for how).

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

Axioms = Normative Requirements

We formulate normative requirements in the form of so-called *axioms*.

Some particularly convincing examples:

- *Participation Principle:* It should be in the best interest of voters to participate; voting truthfully should be no worse than abstaining.
- *Condorcet Principle:* If there is an alternative that is preferred to every other alternative by a majority of voters, then it should win.
- *Pareto Principle:* There should be no alternative that every voter strictly prefers to the alternative selected by the voting rule.

<u>But:</u> surprisingly hard to satisfy! ($\hookrightarrow$)

Plurality with Runoff fails the Participation Principle

No-Show Paradox: Under plurality with runoff, it may be better to abstain than to participate and vote for your favourite alternative!

25 voters: $a \succ b \succ c$
46 voters: $c \succ a \succ b$
24 voters: $b \succ c \succ a$

So b gets eliminated, and then c beats a 70:25 in the runoff.

Now suppose two voters from the first group *abstain:*

23 voters: $a \succ b \succ c$
46 voters: $c \succ a \succ b$
24 voters: $b \succ c \succ a$

Now a gets eliminated, and b beats c 47:46 in the runoff.

P.C. Fishburn and S.J Brams. Paradoxes of Preferential Voting. *Mathematics Magazine*, 1983.

Borda fails the Condorcet Principle

Consider this profile with 11 voters:

4 voters:	$c \succ b \succ a$
3 voters:	$b \succ a \succ c$
2 voters:	$b \succ c \succ a$
2 voters:	$a \succ c \succ b$

Borda elects b, but c is majority-preferred to both a and b.

Preview: We'll get back to this. It's going to get much worse.

Cup Rules fail the Pareto Principle

Rule given by *binary tree*, with the alternatives labelling the leaves.
To progress an alternative needs to *majority*-beat its sibling.

Such *cup rules* may fail the Pareto Principle:

```
        o
       / \
      o   d
     / \
    o   a
   / \
  b   c
```

Consider this profile with three voters:

Ann:	$\mathtt{a} \succ \mathtt{b} \succ \mathtt{c} \succ \mathtt{d}$
Bob:	$\mathtt{b} \succ \mathtt{c} \succ \mathtt{d} \succ \mathtt{a}$
Cindy:	$\mathtt{c} \succ \mathtt{d} \succ \mathtt{a} \succ \mathtt{b}$

d *wins!* (despite being dominated by c)

Exercise: *Do you see how I did this?*

More Axioms: Anonymity and Neutrality

Two really fundamental fairness requirements for a voting rule:

- *Anonymity:* Treat all voters the same.
- *Neutrality:* Treat all alternatives the same.

Exercise: *How would you go about formalising these axioms?*

Resolute and Irresolute Voting Rules

A voting rule is called *resolute* if it returns a single winning alternative for every possible profile of preferences. Very useful to have.

Exercise: *Give an example for a voting rule that is resolute.*

For the rest of this slide only, let us restrict attention to voting rules for scenarios with just *two voters* $(n = 2)$ and *two alternatives* $(m = 2)$.

Exercise: *Show that there exists no resolute voting rule that is "fair" in the sense of being both anonymous and neutral.*

Exercise: *But there still are a couple of irresolute voting rules that are both anonymous and neutral. Give some examples! How do we pick?*

May's Theorem

When there are only *two alternatives*, then all the voting rules we have seen coincide. This is usually called the *simple majority rule* (SMR).

Intuitively, it does the "right" thing. Can we make this precise? *Yes!*

Theorem 1 (May, 1952) *A voting rule for two alternatives satisfies anonymity, neutrality, and positive responsiveness iff it is the SMR.*

Meaning of these *axioms:*

- *anonymity* = voters are treated symmetrically
- *neutrality* = alternatives are treated symmetrically
- *positive responsiveness* = if x is the (sole or tied) winner and one voter switches from y to x, then x becomes the sole winner

Exercise: *One direction is easy. Which one? Prove it!*

K.O. May. A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decisions. *Econometrica*, 1952.

Proof Sketch

We want to prove:

> *A voting rule for two alternatives satisfies anonymity, neutrality, and positive responsiveness iff it is the SMR.*

($\Leftarrow$) Clearly, the simple majority rule has all three properties. ✓

($\Rightarrow$) Assume *#voters* is *odd* (other case: similar) $\leadsto$ no ties

Let a A be the set of voters voting $a \succ b$ and B those voting $b \succ a$. *Anonymity* $\leadsto$ only *number* of ballots of each type matters. Cases:

- If $|A| = |B| + 1$, then *only a wins*. $\leadsto$ By *PR*, a wins for $|A| > |B|$. By *neutrality*, b wins otherwise. So we get the SMR. ✓
- There exist A, B with $|A| = |B|+1$ yet *b wins*. $\leadsto$ Let one a-voter switch to b. By *PR*, now *only* b wins. But now $|B'| = |A'| + 1$, which is symmetric to the first situation, so by *neutrality* a wins. $\lightning$

The Condorcet Jury Theorem

The SMR is not just normatively but also *epistemically* attractive, allowing us to *track the truth* (assuming there is a "correct" choice):

Theorem 2 (Condorcet, 1785) *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.

More Voting Rules: Positional Scoring Rules

We can generalise the idea underlying the Borda rule as follows:

A *positional scoring rule* (PSR) is defined by a so-called *scoring vector* $\boldsymbol{s} = (s_1, \ldots, s_m) \in \mathbb{R}^m$ with $s_1 \geqslant s_2 \geqslant \cdots \geqslant s_m$ and $s_1 > s_m$.

Each voter submits a ranking of the m alternatives. Each alternative receives s_j points for every voter putting it at the jth position.

The alternative(s) with the highest score (sum of points) win(s).

Examples:

- *Borda rule* = PSR with scoring vector $(m-1, m-2, \ldots, 0)$
- *Plurality rule* = PSR with scoring vector $(1, 0, \ldots, 0)$
- *Antiplurality* (or *veto*) *rule* = PSR with scoring vector $(1, \ldots, 1, 0)$
- For any $k < m$, *k-approval* = PSR with $(\underbrace{1, \ldots, 1}_{k}, 0, \ldots, 0)$

Exercise: *Name the rule induced by $\boldsymbol{s} = (9, 7, 5)$. General idea?*

Positional Scoring Rules and the Condorcet Principle

Recall: The Borda rule violates the Condorcet Principle. *Bad scores?*

Consider this example with three alternatives and seven voters:

3 voters: $a \succ b \succ c$

2 voters: $b \succ c \succ a$

1 voter: $b \succ a \succ c$

1 voter: $c \succ a \succ b$

So a is the *Condorcet winner:* a beats both b and c (by 4 to 3).
But any *positional scoring rule* makes b win (because $s_1 \geqslant s_2 \geqslant s_3$):

a: $3 \cdot s_1 + 2 \cdot s_2 + 2 \cdot s_3$

b: $3 \cdot s_1 + 3 \cdot s_2 + 1 \cdot s_3$

c: $1 \cdot s_1 + 2 \cdot s_2 + 4 \cdot s_3$

Thus, *no positional scoring rule* for three (or more) alternatives can possibly satisfy the *Condorcet Principle*.

Dodgson's Rule and its Complexity

Here is a rule that satisfies the Condorcet Principle. It was proposed by C.L. Dodgson (a.k.a. Lewis Carroll, author of *Alice in Wonderland*).

> *If a Condorcet winner exists, elect it. Otherwise, for each alternative x compute the number of adjacent swaps in the individual preferences required for x to become a Condorcet winner. Elect the alternative(s) that minimise that number.*

But this voting rule is particularly hard to compute:

Theorem 3 (Hemaspaandra et al., 1997) *Winner determination for Dodgson's rule is complete for parallel access to NP.*

Writings of C.L. Dodgson. In I. McLean and A. Urken (eds.), *Classics of Social Choice*, University of Michigan Press, 1995.

E. Hemaspaandra, L. Hemaspaandra and J. Rothe. Exact Analysis of Dodgson Elections: Lewis Carroll's 1876 Voting System is Complete for Parallel Access to NP. *Journal of the ACM*, 1997.

Strategic Manipulation

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.

<u>Exercise:</u> *Is there a better voting rule that avoids this problem?*

The Gibbard-Satterthwaite Theorem

Answer to the previous question: *No!* — surprisingly, not only the plurality rule, but *all* "reasonable" rules have this problem.

Theorem 4 (Gibbard-Satterthwaite) *All resolute and surjective voting rules for $\geqslant 3$ alternatives are manipulable or dictatorial.*

Meaning of the terms mentioned in the theorem:

- *resolute* = the rule always returns a single winner (no ties)
- *surjective* = each alternative can win for *some* way of voting
- *dictatorial* = the top alternative of some fixed voter always wins

So this is seriously bad news.

A. Gibbard. Manipulation of Voting Schemes. *Econometrica*, 1973.

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

Complexity as a Barrier against Manipulation

By the Gibbard-Satterthwaite Theorem, any voting rule for $\geqslant 3$ candidates can be manipulated (unless it is dictatorial).

Idea: So it's always *possible* to manipulate, but maybe it's *difficult*!

Theorem 5 (Bartholdi and Orlin, 1991) *The manipulation problem for the rule known as single transferable vote (STV) is NP-complete.*

STV is (roughly) defined as follows:

> *Proceed in rounds. In each round, eliminate the current plurality loser. Stop once only one alternative is left.*

Discussion: NP is a worst-case notion. What about average complexity?

J.J. Bartholdi III and J.B. Orlin. Single Transferable Vote Resists Strategic Voting. *Social Choice and Welfare*, 1991.

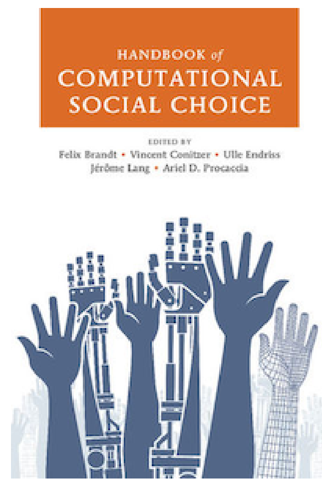
Summary

COMSOC is all about *aggregating* information supplied by *individuals* into a *collective* view. We focused on the *axiomatic* view on *voting:*

- Rules: Plurality, Borda, Scoring Rules, Cup Rules, Dodgson, . . .
- Axioms: Anonymity, Pareto, Condorcet, Strategyproofness, . . .
- Results: May's Theorem, Gibbard-Satterthwaite Theorem, . . .

Tomorrow: *automated reasoning* for social choice and *explainability*

for a much more comprehensive view of the field, consult the *Handbook of COMSOC* (*free download*)



F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A.D. Procaccia (eds), *Handbook of Computational Social Choice*. Cambridge University Press, 2016.