

# Introduction to Computational Social Choice

Ulle Endriss

Institute for Logic, Language and Computation  
University of Amsterdam

[ Guest Lecture for M2R IT Track on Artificial Intelligence  
Université Paul Sabatier, Toulouse ]

## Social Choice Theory

SCT studies collective decision making: how should we aggregate the preferences of the members of a group to obtain a “social preference”?

**Agent 1:**  $\triangle \succ \circ \succ \square$

**Agent 2:**  $\circ \succ \square \succ \triangle$

**Agent 3:**  $\square \succ \triangle \succ \circ$

**Agent 4:**  $\square \succ \triangle \succ \circ$

**Agent 5:**  $\circ \succ \square \succ \triangle$

---

?

SCT is traditionally studied in Economics and Political Science, but now also by “us”: *Computational Social Choice*.

## Social Choice and AI (1)

Social choice theory has natural applications in AI:

- *Multiagent Systems*: to aggregate the beliefs + to coordinate the actions of groups of autonomous software agents
- *Search Engines*: to determine the most important sites based on links (“votes”) + to aggregate the output of several search engines
- *Recommender Systems*: to recommend a product to a user based on earlier ratings by other users
- *AI Competitions*: to determine who has developed *the best* trading agent / SAT solver / RoboCup team

But not all of the classical assumptions will fit these new applications. So AI needs to develop *new models* and *ask new questions*.

## Social Choice and AI (2)

*Vice versa*, techniques from AI, and computational techniques in general, are useful for advancing the state of the art in social choice:

- *Algorithms and Complexity*: to develop algorithms for (complex) voting procedures + to understand the hardness of “using” them
- *Knowledge Representation*: to compactly represent the preferences of individual agents over large spaces of alternatives
- *Logic and Automated Reasoning*: to formally model problems in social choice + to automatically verify (or discover) theorems

Indeed, you will find many papers on social choice at AI conferences (e.g., IJCAI, ECAI, AAI, AAMAS, KR) and many AI researchers participate in events dedicated to social choice (e.g., COMSOC).

F. Brandt, V. Conitzer, and U. Endriss. Computational Social Choice. In G. Weiss (ed.), *Multiagent Systems*, MIT Press, 2013.

## Plan for the Remainder of this Lecture

The purpose of today's lecture is to provide you with both:

- an overview of types of collective decision making *problems*, and
- an overview of *techniques* used to address these problems.

These are the types of problems we will consider:

- *fair allocation of goods*: e.g., computing resources to users
- *two-sided matching*: e.g., junior doctors to hospitals
- *voting*: e.g., for candidates in political elections
- *judgment aggregation*: e.g., regarding annotated data in linguistics

## Fair Allocation of Goods

Consider a set of agents and a set of goods. Each agent has her own preferences regarding the allocation of goods to agents. Examples:

- allocation of resources amongst members of our society
- allocation of bandwidth to processes in a communication network
- allocation of compute time to scientists on a super-computer
- ...

We will focus on one specific model studied in the literature, with a single good that can be divided into arbitrarily small pieces ...



## Cut-and-Choose

The classical approach for dividing a cake between *two agents*:

- ▶ One agent *cuts* the cake in two pieces (she considers to be of equal value), and the other *chooses* one of them (the piece she prefers).

The cut-and-choose protocol is *fair* in the sense of guaranteeing a property known as *proportionality*:

- Each agent is *guaranteed* at least one half (general:  $1/n$ ), according to her own valuation.
- Discussion: In fact, the first agent (if she is risk-averse) will receive exactly  $1/2$ , while the second will usually get more.

What if there are *more than two* agents?



## The Banach-Knaster Last-Diminisher Protocol

In the first ever paper on fair division, Steinhaus (1948) reports on a *proportional* protocol for *n agents* due to Banach and Knaster.

- (1) Agent 1 cuts off a piece (that she considers to represent  $1/n$ ).
- (2) That piece is passed around the agents. Each agent either lets it pass (if she considers it too small) or trims it down further (to what she considers  $1/n$ ).
- (3) After the piece has made the full round, the last agent to cut something off (the “last diminisher”) is obliged to take it.
- (4) The rest (including the trimmings) is then divided amongst the remaining  $n-1$  agents. Play cut-and-choose once  $n = 2$ . ✓

Each agent is guaranteed a *proportional* piece. Requires  $O(n^2)$  cuts. May not be *contiguous* (unless you always trim “from the right”).

H. Steinhaus. The Problem of Fair Division. *Econometrica*, 16:101–104, 1948.

## The Even-Paz Divide-and-Conquer Protocol

Even and Paz (1984) introduced the divide-and-conquer protocol:

- (1) Ask each agent to put a *mark* on the cake.
- (2) *Cut* the cake at the  $\lfloor \frac{n}{2} \rfloor$ th *mark* (counting from the left).  
Associate the agents who made the *leftmost*  $\lfloor \frac{n}{2} \rfloor$  *marks* with the *lefthand part*, and the *remaining agents* with the *righthand part*.
- (3) *Repeat* for each group, until only one agent is left.

This also is *proportionally fair*.

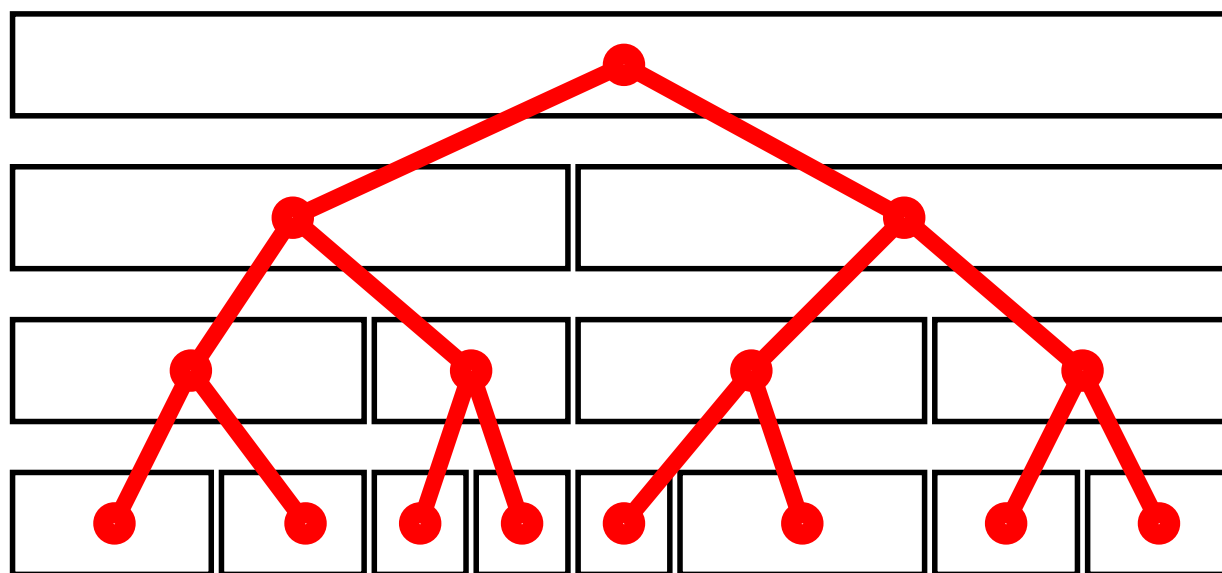
Exercise: How *complex* is this (how many marks)?

S. Even and A. Paz. A Note on Cake Cutting. *Discrete Applied Mathematics*, 7(3):285–296, 1984.

## Complexity Analysis: Number of Marks

In each round, every agent makes one mark. So:  *$n$  marks per round*

But how many rounds?



*rounds* = number of times you can divide  $n$  by 2 before hitting  $\leq 1$   
 $\approx \log_2 n$  (example:  $\log_2 8 = 3$ )

Thus: the number of *marks* is  $O(n \cdot \log n)$ , i.e., much better than for the last-diminisher protocol.

## Preferences

For the cake-cutting scenario, we made some very specific assumptions regarding the *preferences* of the agents:

- preferences are modelled as *utility functions*
- those preferences are *additive* (severe restriction)

Discussion: *cardinal* utility function vs. *ordinal* preference relation

We also did not worry about what formal *language* to use to *represent* an agent's preferences, e.g., to be able to say *how much information* you need to exchange when eliciting an agent's preferences.

Preference representation is an interesting field in its own right.

A possible starting point is the survey cited below.

Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. Preference Handling in Combinatorial Domains: From AI to Social Choice. *AI Magazine*, 29(4):37–46, 2008.

## Matching

In a variant of the fair allocation problem, we try to match each agent with a single item—which may have preferences itself. Examples:

- children to schools
- junior doctors to hospitals
- kidney patients to kidney donors
- ...

We now briefly look into *the* classical matching problem.

## The Stable Marriage Problem

We are given:

- $n$  *men* and  $n$  *women*
- each has a linear *preference* ordering over the opposite sex

We seek:

- a *stable* matching of men to women: no man and woman should want to divorce their assigned partners and run off with each other

## The Gale-Shapley Algorithm

**Theorem 1 (Gale and Shapley, 1962)** *There exists a stable matching for any combination of preferences of men and women.*

The *Gale-Shapley “deferred acceptance” algorithm* for computing a stable matching works as follows:

- In each round, each man who is not yet engaged proposes to his favourite amongst the women he has not yet proposed to.
- In each round, each woman picks her favourite from the proposals she’s receiving and the man she’s currently engaged to (if any).
- Stop when everyone is engaged.

D. Gale and L.S. Shapley. College Admissions and the Stability of Marriage. *American Mathematical Monthly*, 69:9–15, 1962.

## Voting

In voting theory, each agent is assumed to have a linear preference order over a set of alternatives, and based on this information we want to elect the “best” alternative. Examples:

- voting in a political election
- aggregating advice received from several experts
- decision making in a multiagent system



## Three Voting Rules

In voting,  $n$  *voters* choose from a set of  $m$  *alternatives* by stating their preferences in the form of *linear orders* over the alternatives.

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

- *Plurality*: elect the alternative ranked first most often (i.e., each voter assigns 1 point to an alternative of her choice, and the alternative receiving the most points wins)
- *Plurality with runoff*: run a plurality election and retain the two front-runners; then run a majority contest between them
- *Borda*: each voter gives  $m-1$  points to the alternative she ranks first,  $m-2$  to the alternative she ranks second, etc.; and the alternative with the most points wins

## Example: Choosing a Beverage for Lunch

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

2 *Germans*: Beer  $\succ$  Wine  $\succ$  Milk  
3 *Frenchmen*: Wine  $\succ$  Beer  $\succ$  Milk  
4 *Dutchmen*: Milk  $\succ$  Beer  $\succ$  Wine

Which beverage *wins* the election for

- the plurality rule?
- plurality with runoff?
- the Borda rule?

## Axiomatic Method

So how do you decide which is the right voting rule to use?

The classical approach is to use the *axiomatic method*:

- identify good axioms: normatively appealing high-level properties
- give mathematically rigorous definitions of these axioms
- explore the consequences of the axioms

The definitions on the following slide are only sketched, but can be made mathematically precise (see the 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.

## 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 2 (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  $A$  is the (sole or tied) winner and one voter switches from  $B$  to  $A$ , then  $A$  becomes the sole winner

K.O. May. A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decisions. *Econometrica*, 20(4):680–684, 1952.

## Proof Sketch

We want to prove:

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

Proof: Clearly, the simple majority rule has all three properties. ✓

Other direction: assume *#voters* is *odd* (other case: similar)  $\leadsto$  no ties

Let a  $\mathcal{A}$  be the set of voters voting  $A \succ B$  and  $\mathcal{B}$  those voting  $B \succ A$ .

*Anonymity*  $\leadsto$  only number of ballots of each type matters. Two cases:

- Whenever  $|\mathcal{A}| = |\mathcal{B}| + 1$  then *only A wins*. Then, by *PR*,  $A$  wins whenever  $|\mathcal{A}| > |\mathcal{B}|$  (which is exactly the simple majority rule). ✓
- There exist  $\mathcal{A}, \mathcal{B}$  with  $|\mathcal{A}| = |\mathcal{B}| + 1$  but *B wins*. Let one  $A$ -voter switch to  $B$ . By *PR*, now only  $B$  wins. But now  $|\mathcal{B}'| = |\mathcal{A}'| + 1$ , which is symmetric to the first situation, so by *neutrality*  $A$  wins. ✗

## The Condorcet Jury Theorem

The simple majority rule for two alternatives is attractive also in terms of *truth-tracking* (assuming there is a “correct” choice):

**Theorem 3 (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  $p \cdot n > \frac{1}{2} \cdot n$ . ✓

For a modern exposition, see Young (1995).

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

H.P. Young. Optimal Voting Rules. *J. Economic Perspectives*, 9(1):51–64, 1995.

## Positional Scoring Rules

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

A *positional scoring rule* is given by a *scoring vector*  $s = \langle s_1, \dots, s_m \rangle$  with  $s_1 \geq s_2 \geq \dots \geq s_m$  and  $s_1 > s_m$ .

Each voter submits a ranking of the  $m$  alternatives. Each alternative receives  $s_i$  points for every voter putting it at the  $i$ th position.

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

Examples:

- *Borda rule* = PSR with scoring vector  $\langle m-1, m-2, \dots, 0 \rangle$
- *Plurality rule* = PSR with scoring vector  $\langle 1, 0, \dots, 0 \rangle$
- *Anti-plurality rule* = PSR with scoring vector  $\langle 1, \dots, 1, 0 \rangle$
- For any  $k \leq m$ , *k-approval* = PSR with  $\underbrace{\langle 1, \dots, 1, 0, \dots, 0 \rangle}_k$

## The Condorcet Principle

Another idea going back to Condorcet: an alternative beating all other alternatives in pairwise majority contests is a *Condorcet winner*.

Sometimes there is no Condorcet winner (*Condorcet paradox*):

Ann:  $A \succ B \succ C$

Bob:  $B \succ C \succ A$

Cindy:  $C \succ A \succ B$

But if a Condorcet winner exists, then it must be *unique*.

A voting rule satisfies the *Condorcet principle*, if it elects (only) the Condorcet winner whenever one exists.



## All PSR's Violate Condorcet

Consider the following example:

3 voters:  $A \succ B \succ C$

2 voters:  $B \succ C \succ A$

1 voter:  $B \succ A \succ C$

1 voter:  $C \succ A \succ B$

$A$  is the *Condorcet winner*; she beats both  $B$  and  $C$  4 : 3. But any *positional scoring rule* makes  $B$  win (because  $s_1 \geq s_2 \geq s_3$ ):

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

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

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

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 4 (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*, 44(6):806–825, 1997.

## Example: Strategic Manipulation

Suppose the *plurality rule* is used to decide an election: the candidate ranked first most often wins. Recall Florida in 2000 (simplified):

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

Bush will win this election. It would have been in the interest of the Nader supporters to pretend that they like Gore the most.

Thus, the plurality is subject to *strategic manipulation*: sometimes, some voters can get a better outcome by lying about their preferences.

► 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 5 (Gibbard-Satterthwaite)** *All resolute and surjective voting rules for  $\geq 3$  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: A General Result. *Econometrica*, 41(4):587–601, 1973.

M.A. Satterthwaite. Strategy-proofness and Arrow’s Conditions. *Journal of Economic Theory*, 10:187–217, 1975.

## Logic for Social Choice Theory

Nowadays, the (omitted) proof of the Gibbard-Satterthwaite Theorem is well understood, but after people developed good intuitions that something like G-S must be the case in the 1960's, it still took around a decade before someone was able to prove it. So this is not trivial!

Idea: Cast this in a suitable *logic* and use *automated theorem provers*!

Indeed, this works to some extent (but is still an underdeveloped area):

- Nipkow (2009) *verified* a known proof for G-S in ISABELLE.
- For related results, proofs have also been *derived* automatically, and some simpler results even have been *discovered* automatically.

T. Nipkow. Social Choice Theory in HOL. *J. Autom. Reas.*, 43(3):289–304, 2009.

P. Tang and F. Lin. Computer-aided Proofs of Arrow's and other Impossibility Theorems. *Artificial Intelligence*, 173(11):1041–1053, 2009.

C. Geist and U. Endriss. Automated Search for Impossibility Theorems in Social Choice Theory: Ranking Sets of Objects. *J. Artif. Intell. Res.*, 40:143-174, 2011.

## Complexity as a Barrier against Manipulation

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

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

**Theorem 6 (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.*

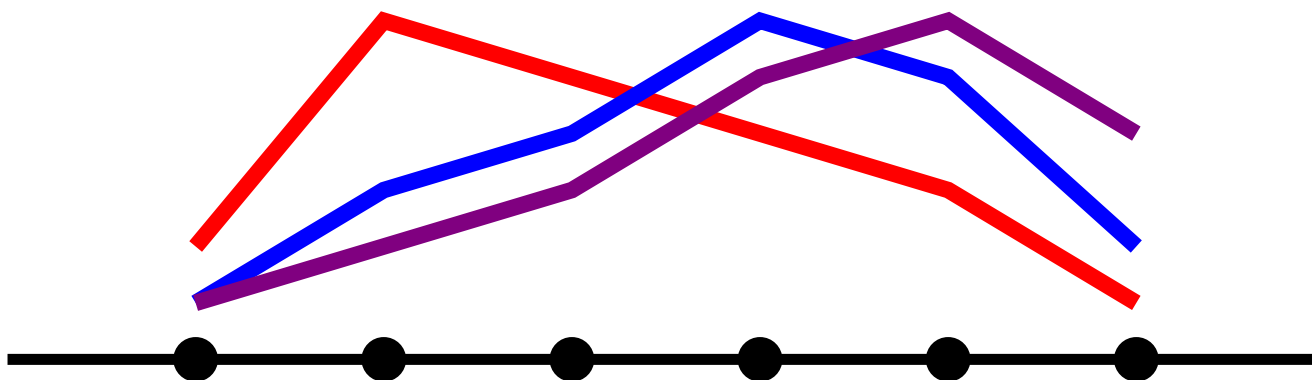
Discussion: NP is a worst-case notion; what about average complexity?

J.J. Bartholdi III and J.B. Orlin. Single Transferable Vote Resists Strategic Voting. *Social Choice and Welfare*, 8(4):341–354, 1991.

## Domain Restrictions

The G-S Theorem applies to voting rules that need to work for *all* preference profiles. But they often come with some inherent structure.

For example, sometimes preferences are *single-peaked* with respect to some natural left-to-right ordering of the alternatives:



Good news: now the *median-voter rule* is immune to manipulation!

D. Black. On the Rationale of Group Decision-Making. *The Journal of Political Economy*, 56(1):23–34, 1948.

## Social Choice in Combinatorial Domains

Suppose 13 voters are asked to each vote *yes* or *no* on three issues; and we use the simple majority rule for each issue independently:

- 3 voters each vote for YNN, NYN, NNY.
- 1 voter each votes for YYY, YYN, YNY, NYY.
- No voter votes for NNN.

But then NNN wins: on each issue, 7 out of 13 vote *no* (*paradox!*)

*What to do instead?* The number of candidates is *exponential* in the number of issues (e.g.,  $2^3 = 8$ ), so even just representing the voters' preferences is a challenge ( $\rightsquigarrow$  *knowledge representation*).

S.J. Brams, D.M. Kilgour, and W.S. Zwicker. The Paradox of Multiple Elections. *Social Choice and Welfare*, 15(2):211–236, 1998.

Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. Preference Handling in Combinatorial Domains: From AI to Social Choice. *AI Magazine*, 29(4):37–46, 2008.



## Judgment Aggregation

Preferences are not the only structures we may wish to aggregate. We can also aggregate judgments, opinions, beliefs. Examples:

- judges collectively presiding over a trial
- robots integrating their individual sensor information
- a scientist aggregating data crowdsourced from multiple sources

U. Endriss. Judgment Aggregation. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia (eds.), *Handbook of Computational Social Choice*, CUP, 2016.

## Example

Suppose three robots are in charge of climate control for this building. They need to make judgments on  $p$  (the temperature is below  $17^\circ\text{C}$ ),  $q$  (we should switch on the heating), and  $p \rightarrow q$ .

	$p$	$p \rightarrow q$	$q$
Robot 1:	Yes	Yes	Yes
Robot 2:	No	Yes	No
Robot 3:	Yes	No	No

► What should be the collective decision?

## Summary

COMSOC is all about *aggregating* information supplied by *individuals* into a *collective* view. Different *domains* of aggregation:

- *fair allocation*: preferences over highly structured alternatives
- *matching*: two groups of agents with preferences over each other
- *voting*: ordinal preferences over alternatives w/o internal structure
- *judgment aggregation*: assignments of truth values to propositions

Different *techniques* used to analyse them, such as:

- axiomatic method: philosophical and mathematical
- logical modelling, automated theorem proving
- algorithm design and complexity analysis
- probability theory (e.g., for truth-tracking)

For more information, see the website of my Amsterdam course:

<https://staff.science.uva.nl/u.endriss/teaching/comsoc/>