Computational Social Choice:  
Guest Lecture for Autonomous Agents, Dec. 2014

Ulle Endriss
Institute for Logic, Language and Computation
University of Amsterdam
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: \( \triangle \succ \bigcirc \succ \square \)
Agent 2: \( \bigcirc \succ \square \succ \triangle \)
Agent 3: \( \square \succ \triangle \succ \bigcirc \)
Agent 4: \( \square \succ \triangle \succ \bigcirc \)
Agent 5: \( \bigcirc \succ \square \succ \triangle \)

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

- Why social choice and AI?
- Some voting rules and an example
- Main topic for today: Strategic voting
- Research topics in 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, AAAI, AAMAS) and many AI researchers participate in events dedicated to social choice (e.g., COMSOC).

Three Voting Rules

How should $n$ voters choose from a set of $m$ 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 ≻ Wine ≻ Milk
3 Frenchmen: Wine ≻ Beer ≻ Milk
4 Dutchmen: Milk ≻ Beer ≻ Wine

Which beverage wins the election for

- the plurality rule?
- plurality with runoff?
- the Borda rule?
Example: Electing a President

Remember Florida 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

Questions:

- Who wins?
- Is that a fair outcome?
- What would your advice to the Nader-supporters have been?
Notation and Terminology

Finite set of $n$ voters (or individuals or agents) $\mathcal{N} = \{1, \ldots, n\}$.

Finite set of $m$ alternatives (or candidates) $\mathcal{X}$.

Each voter expresses a preference over the alternatives by providing a linear order on $\mathcal{X}$ (her ballot). $\mathcal{L}(\mathcal{X})$ is the set of all such linear orders.

A profile $\mathbf{R} = (R_1, \ldots, R_n)$ fixes one preference/ballot for each voter.

A (resolute) voting rule or is a function $F$ mapping any given profile to a (single) winning alternative:

$$F : \mathcal{L}(\mathcal{X})^n \rightarrow \mathcal{X}$$
**Strategy-Proofness**

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

\(F\) is *strategy-proof* (or *immune to manipulation*) if for no individual \(i \in \mathcal{N}\) there exist a profile \(R\) (including the “truthful preference” \(R_i\) of \(i\)) and a linear order \(R'_i\) (representing the “untruthful” ballot of \(i\)) such that \(F(R_{-i}, R'_i)\) is ranked above \(F(R)\) according to \(R_i\).

In other words: under a strategy-proof voting rule no voter will ever have an incentive to misrepresent her preferences.

*Note:* Strategy-proofness is a very desirable property! We have seen that *plurality is not strategy-proof*. From the lecture on mechanism design, you know that, e.g., *strategy-proof auction mechanisms* do exist.
The Gibbard-Satterthwaite Theorem

Two more properties of resolute voting rules $F$:

- $F$ is **surjective** if for every candidate $x \in \mathcal{X}$ there exists a profile $R$ such that $F(R) = x$.

- $F$ is a **dictatorship** if there exists a voter $i \in \mathcal{N}$ (the dictator) such that $F(R) = \text{top}(R_i)$ for any profile $R$.

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

**Theorem 1 (Gibbard-Satterthwaite)** Any resolute voting rule for $\geq 3$ candidates that is surjective and strategy-proof is a dictatorship.

---


Remarks

The G-S Theorem says that for $\geq 3$ candidates, any resolute voting rule $F$ that is surjective and strategy-proof is a dictatorship.

- a surprising result + not applicable in case of two candidates
- The opposite direction is clear: dictatorial $\Rightarrow$ strategy-proof
- Random procedures don’t count (but might be “strategy-proof”).

We will now prove the theorem under two additional assumptions:

- $F$ is neutral, i.e., candidates are treated symmetrically.
  [Note: neutrality $\Rightarrow$ surjectivity; so we won’t make use of surjectivity.]
- There are exactly 3 candidates.

For a full proof, using a similar approach, see, e.g.:

Proof (1): Independence and Blocking Coalitions

Notation: $N^R_{x>y}$ is the set of voters who rank $x$ above $y$ in profile $R$.

Claim: If $F(R) = x$ and $N^R_{x>y} = N^{R'}_{x>y}$, then $F(R') \neq y$. [independence]

Proof: From strategy-proofness, by contradiction. Assume $F(R') = y$. Moving from $R$ to $R'$, there must be a first voter to affect the winner. So w.l.o.g., assume $R$ and $R'$ differ only w.r.t. voter $i$. Two cases:

- $i \in N^R_{x>y}$: Suppose $i$’s true preferences are as in profile $R'$ (i.e., $i$ prefers $x$ to $y$). Then $i$ has an incentive to vote as in $R$. ✓
- $i \notin N^R_{x>y}$: Suppose $i$’s true preferences are as in profile $R$ (i.e., $i$ prefers $y$ to $x$). Then $i$ has an incentive to vote as in $R'$. ✓

Some more terminology:

Call $C \subseteq N$ a blocking coalition for $(x, y)$ if $C = N^R_{x>y} \Rightarrow F(R) \neq y$.

Thus: If $F(R) = x$, then $C := N^R_{x>y}$ is blocking for $(x, y)$ [for any $y$].
Proof (2): Ultrafilters

From *neutrality*: all \((x, y)\) must have *the same* blocking coalitions.

For any \(C \subseteq \mathcal{N}\), \(C\) or \(\overline{C} := \mathcal{N} \setminus C\) must be blocking.

Proof: Assume \(C\) is not blocking; i.e., \(C\) is not blocking for \((x, y)\).
Then there exists a profile \(R\) with \(N^R_{x \succ y} = C\) but \(F(R) = y\).
But we also have \(N^R_{y \succ x} = \overline{C}\). Hence, \(\overline{C}\) is blocking for \((y, x)\).

If \(C_1\) and \(C_2\) are blocking, then so is \(C_1 \cap C_2\). [now we'll use \(|\mathcal{X}| = 3\)]

Proof: Consider a profile \(R\) with \(C_1 = N^R_{x \succ y}\), \(C_2 = N^R_{y \succ z}\), and \(C_1 \cap C_2 = N^R_{x \succ z}\). As \(C_1\) is blocking, \(y\) cannot win. As \(C_2\) is
blocking, \(z\) cannot win. So \(x\) wins and \(C_1 \cap C_2\) must be blocking.

The *empty coalition* is *not* blocking.

Proof: Omitted (but not at all surprising).

Above properties (+ finiteness of \(\mathcal{N}\)) imply that there’s a *singleton* \(\{i\}\) that is blocking. But that just means that \(i\) is a *dictator*! ✓
Relationship to Mechanism Design

So why does this work for auctions and not for voting?

Recall that the Vickrey Auction is strategy-proof:

One item on auction. Each bidder makes a bid (offering a price).
The highest bidder wins, but pays the second highest price. No incentive to misrepresent true price.

We could think of this as an election:

- Candidates = set of all pairs of the form \((\text{winner}, \text{price})\).
- Preference order determined by bid. E.g., if \textit{Alice} has valuation \(€5\):

\[
\cdots \succ (\text{Alice}, 3) \succ (\text{Alice}, 4) \succ (\text{Alice}, 5) \sim (Bob, 17) \sim \cdots \\
\cdots \sim (Carla, 2) \succ \cdots \succ (\text{Alice}, 6) \succ (\text{Alice}, 7) \succ \cdots
\]

That is: very specific preferences ("quasi-linear utilities")

This domain restriction is the underlying reason why we can turn the general impossibility (Gibbard-Satterthwaite) into a specific possibility (Vickrey).
Single-Peakedness

The G-S Thm shows that no “reasonable” voting rule is strategy-proof.

Another important domain restriction is due to Black (1948):

- **Definition:** A profile is *single-peaked* if there exists a “left-to-right” ordering $\gg$ on the candidates such that any voter ranks $x$ above $y$ if $x$ is between $y$ and her top candidate w.r.t. $\gg$. Think of spectrum of political parties.

- **Result:** Fix a dimension $\gg$. Assuming that all profiles are single-peaked w.r.t. $\gg$, the *median-voter rule* is strategy-proof.

**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!

Tools from complexity theory can be used to make this idea precise.

- For some procedures this does not work: if I know all other ballots and want $X$ to win, it is easy to compute my best strategy.
- But for others it does work: manipulation is NP-complete.

Recent work in COMSOC has expanded on this idea:

- NP is a worst-case notion. What about average complexity?
- Also: complexity of winner determination, control, bribery, . . .


Automated Reasoning for Social Choice Theory

Logic has long been used to formally specify computer systems, facilitating verification of properties. Can we apply this methodology also here? Yes:

- Verification of a (known) proof of the Gibbard-Satterthwaite Theorem in the HOL proof assistant Isabelle (Nipkow, 2009).
- Fully automated proof of Arrow’s Theorem for 3 candidates via a SAT solver or constraint programming (Tang and Lin, 2009).
- Automated search for new impossibility theorems in ranking sets of objects using a SAT solver (Geist and Endriss, 2011).


Social Choice in Combinatorial Domains

Suppose 13 voters are asked to each vote yes or no on three issues; and we use the plurality 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: 7 out of 13 vote no on each issue (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 (knowledge representation).


This has been an introduction to *computational social choice*: the study of collective decision making using computational (and other) methods.

We have focussed on *strategic behaviour* in *elections*:

- Gibbard-Satterthwaite Thm: no voting rules can prevent manipulation
- But ok for certain domain restrictions: quasi-linearity, single-peakedness
- Computational complexity might also provide protection

COMSOC is a fast moving research area with many opportunities for AI.

To find out more:

- Read our chapter in the MAS book (it's on my website)
- Take my course (starts in February):
  
  http://www.illc.uva.nl/~ulle/teaching/comsoc/