Logic and Social Choice Theory

Ulle Endriss

Institute for Logic, Language and Computation

University of Amsterdam

$$\left[\begin{array}{c} \text{Guest Lecture for ``Logic, Language and Computation''} \\ \text{Master of Logic, 8 November 2019} \end{array} \right]$$

Social Choice Theory

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* to the *algorithmic* (the latter in *computational social choice*).

Its findings are relevant to a variety of *applications*, such as these:

- How to choose a president given people's preferences?
- How to combine the website rankings of multiple search engines?
- How to fairly divide resources between several stake-holders?
- How to assign student doctors to hospitals?
- How to aggregate the views of different judges in a court case?

The most widely studied scenario is that of *preference aggregation*, with n agents each ranking m alternatives.

Outline

- Examples for preference aggregation scenarios and rules
- Example for a classical result: Arrow's Theorem (with proof)
- Examples for recent (mostly local) work on logic and SCT

The Plurality Rule

Under the *plurality rule*, the alternative ranked first most often wins:

$$
\begin{array}{ll}
\text{Voter 1:} & x \succ y \succ w \succ z \\
\text{Voter 2:} & x \succ y \succ w \succ z \\
\text{Voter 3:} & y \succ z \succ w \succ x \\
\text{Voter 4:} & z \succ y \succ w \succ x
\end{array}
$$

The most widely used rule in practice (by far!). *But is it any good?*

The Borda Rule

Under the *Borda rule*, each voter gives $m-1$ points to the alternative she ranks first, $m-2$ to the alternative she ranks second, and so forth:

Voter 1:	$x \succ z \succ y$	$x : 2 + 2 + 1 + 1 + 0 = 6$
Voter 2:	$x \succ z \succ y$	$y : 0 + 0 + 2 + 2 + 1 = 5$
Voter 3:	$y \succ x \succ z$	$z : 1 + 1 + 0 + 0 + 2 = 4$
Voter 4:	$y \succ x \succ z$	
Voter 5:	$z \succ y \succ x$	

A clear advantage over the plurality rule is that we use much more of the information present in the profile to come to a decision.

But there still is a problem: *What if y challenges the winner x?*

The Condorcet Rule

Under the *Condorcet rule*, we run one-to-one majority contests between all pairs of alternatives and elect the one that performs best.

Nice idea. But: *Do you see the problem with this "definition"?*

$$\begin{array}{ll}
\text{Voter 1:} & x \succ y \succ z \\
\text{Voter 2:} & y \succ z \succ x \\
\text{Voter 3:} & z \succ x \succ y
\end{array}$$

The Axiomatic Method

So how do you decide what is the right aggregation 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 your axioms
- explore the logical consequences of your definitions

The Model

The *voters* from a finite set $N = \{1, \ldots, n\}$, with $n \geqslant 2$, all rank the *alternatives* in a set X by supplying a *strict linear order* in $\mathcal{L}(X)$.

We are interested in *preference aggregation rules* of this form:

$$F : \mathcal{L}(X)^n \to \mathcal{L}(X)$$

<u>Thus:</u> Given a *profile* $(\succ_1, \ldots, \succ_n) \in \mathcal{L}(X)^n$ of rankings, one for each voter, we want our rule to return a single collective ranking.

<u>Remark:</u> Alternatively, we could study rules $F : \mathcal{L}(X)^n \to X$ to return the "best" alternative. Essentially the same. *Do you see why?*

First Axiom: The Pareto Condition

A preference aggregation rule F satisfies the (weak) *Pareto condition* if it ranks x above y whenever all individual voters do:

$$\{i \in N \mid x \succ_i y\} = N \quad \Rightarrow \quad x \succ y \ \text{ for } \succ = F(\succ_1, \ldots, \succ_n)$$

Discussion: *Desirable to use rules that satisfy this axiom?*

Second Axiom: Arrow's Independence Condition

An aggregation rule F satisfies Arrow's condition of the *independence of irrelevant alternatives* (IIA) if the relative ranking of x and y it returns depends only on the relative rankings of x and y in the profile:

$$\{i \mid x \succ_i y\} = \{\, i \mid x \succ'_i y\} \;\Rightarrow\; F(\succ_1, \ldots, \succ_n) =_{x?y} F(\succ'_1, \ldots, \succ'_n)$$

<u>Discussion:</u> *Desirable to use rules that satisfy this axiom?*

Summary of Desiderata

We would like to find a rule $F : \mathcal{L}(X)^n \to \mathcal{L}(X)$ that satisfies both the Pareto condition an IIA. *How about the rules discussed so far?*

- Plurality
- Borda
- Condorcet

Remark: These rules were originally formulated for finding a winning alternative (not a ranking) and all of them might run into tie-breaking issues. We can handle these. Focus on the bigger picture instead.

Arrow's Impossibility Theorem

A rule F is a *dictatorship* if there exists a "dictator" $i^\star \in N$ such that $F(\succ_1, \ldots, \succ_n) = \succ_{i^\star}$ for every profile $(\succ_1, \ldots, \succ_n)$. *Bad!*

Theorem 1 (Arrow, 1951) *Any aggregation rule for* $\geqslant 3$ *alternatives that satisfies the Pareto condition and IIA must be a dictatorship.*

Remarks:

- not true for 2 alternatives (*counterexample?*)
- opposite direction also holds (so this is a characterisation result)
- dictatorial does *not* just mean: outcome = someone's preference

Next: Proof (following Geanakoplos, 2005).

K.J. Arrow. *Social Choice and Individual Values*. John Wiley and Sons, 2nd edition, 1963. First edition published in 1951.

J. Geanakoplos. Three Brief Proofs of Arrow's Impossibility Theorem. *Economic Theory*, 26(1):211–215, 2005.

Extremal Lemma

Assume there are $\geqslant 3$ alternatives and F satisfies Pareto and IIA.

<u>Claim:</u> If all individuals rank y either top or bottom, then so does F.

<u>Proof:</u> Suppose otherwise, i.e., all individuals rank alternative y either top or bottom, but F does not. Write $\succ$ for the ranking returned by F.

(1) Then there exist alternatives $x, z \in X$ such that $x \succ y$ and $y \succ z$.

(2) By *IIA*, this does not change if we *move z above x* in every individual ranking (as doing so we don't cross the extremal y).

(3) By *Pareto*, in the new profile we must have $z \succ x$.

(4) But we still have $x \succ y$ and $y \succ z$, so by *transitivity* we get $x \succ z$.

Contradiction. ✓

Existence of an Extremal-Pivotal Voter

Fix some alternative y. Call a voter *extremal-pivotal* if she can push y from the bottom to the top in F for at least one profile.

Claim: There exists an extremal-pivotal voter (for y).

Proof: Consider a profile where every voter ranks y *at the bottom*. By Pareto, so does F. Let voters *switch y to the top*, one by one. By the Extremal Lemma, after each step, y *is still extremal* in F. By Pareto, at the end of this process, F ranks y *at the top*.

So there must be a point where y jumps from the bottom to the top. The individual making the corresponding switch is extremal-pivotal. ✓

Remark: Only works because N is finite. *Do you see why?*

Notation: Let α the profile just before the jump and β the profile just after the jump. Let i be the extremal-pivotal voter we just found.

Dictatorship for Most Alternatives

<u>Recall</u>: i is extremal-pivotal for y in α (y is at bottom for i and F),
from where she can force β (y is at top for i and F).

<u>Claim</u>: Voter i *can dictate* the relative ranking under F of any two
alternatives x *and* z that are different from y.

<u>Proof</u>: W.l.o.g., suppose i wants to place x above z.

Let i vote almost as in β, but moving x to the top: $x \succ_i y \succ_i z$.
Let all others rank y *as in* β, but otherwise (including the relative
ranking of x and z) vote as they wish. Consider the resulting profile γ:

- Observe that in γ all relative rankings of x and y are as in α.
  So by *IIA*, we must still have $x \succ y$.
- Observe that in γ all relative rankings of y and z are as in β.
  So by *IIA*, we must still have $y \succ z$.

By *transitivity*, we get $x \succ z$. By *IIA*, this continues to hold if others
change their relative rankings of alternatives other than x and z. ✓

Dictatorship for All Alternatives

Recall: Voter i can dictate relative rankings of x and z (anything $\neq y$).

Claim: Voter i can also dictate relative rankings involving y.

Proof: We can use a similar construction as before to show:

- Some voter j can dictate relative rankings of x and y ($\neq z$).
- Some voter k can dictate relative rankings of y and z ($\neq x$).

But then, exploiting *transitivity*, j and k together can dictate relative rankings of x and z. As there cannot be two different dictators on the same pair, voters i, j and k must all be the same individual. ✓

Thus, voter i is a dictator for any two alternatives, meaning that the rule F is dictatorial. This proves Arrow's Theorem.

Logic and Social Choice Theory

The axiomatic method itself is an (informal) use of logic in SCT.

<u>Next:</u> Several more explicit uses of logic in research on SCT . . .

Formal Modelling of Social Choice Scenarios

Maybe the most obvious application of logic concerns the modelling of social choice scenarios using an appropriate logical language:

- 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)

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

Logic has long been used to help verify the correctness of hardware and software. Can we use this methodology also here? *Yes!*

- Automated *verification* of a (known) proof of Arrow's Theorem in the HOL proof assistant ISABELLE (Nipkow, 2009).

- Automated *proof* of Arrow's Theorem for 3 alternatives and 2 voters using a SAT-solver (Tang and Lin, 2009).

- Automated *discovery* of new impossibility theorems in *ranking sets of objects* using a SAT-solver (Geist and Endriss, 2011).

T. Nipkow. Social Choice Theory in HOL. *J. Autom. Reas.*, 2009.

P. Tang and F. Lin. Computer-aided Proofs of Arrow's and other Impossibility Theorems. *Artificial Intelligence*, 2009.

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

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 alternatives is *exponential* in the number of issues (e.g., $2^3 = 8$), so even just representing the voters' preferences is a challenge ($\hookrightarrow$ *logic-based knowledge representation*).

S.J. Brams, D.M. Kilgour, and W.S. Zwicker. The Paradox of Multiple Elections. *Social Choice and Welfare*, 1998.

Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. Preference Handling in Combinatorial Domains: From AI to Social Choice. *AI Magazine*, 2008.

Judgment Aggregation

Preferences are not the only structures you may wish to aggregate . . .

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

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

What to do?

C. List and P. Pettit. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy*, 2002.

U. Endriss. Judgment Aggregation. In Brandt et al. (eds), *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

Last Slide

We have seen:

- Arrow's Theorem as an example for classical work in social choice theory: there can be no perfect preference aggregation rule.

- Examples for modern research directions in computational social choice (emphasising logic in particular).

For possible entry points into the field, see the papers cited below, my MoL course on COMSOC, or this website:

$$\texttt{http://research.illc.uva.nl/COMSOC/}$$

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

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