Introduction to Computational Social Choice

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

Lecture 1

Social choice theory studies mechanisms for collective decision making, such as voting procedures or fair division protocols.

This course will be an introduction to *computational social choice*, presenting both classical work in social choice theory and highlighting computational aspects and applications of logic.

Outline of this first lecture:

- A couple of introductory examples
- Arrow's Theorem, *the* classical result in social choice theory
- Overview of the COMSOC research area
- Outlook on what will happen during the rest of the course

Example from Voting

Suppose the *plurality rule* (as in most real-world situations) is used to decide the outcome of an election: the candidate receiving the highest number of votes wins.

Assume the preferences of the people in, say, Florida are as follows:

$$\begin{array}{ll} 49\%: & \text{Bush} \succ \text{Gore} \succ \text{Nader} \\ 20\%: & \text{Gore} \succ \text{Nader} \succ \text{Bush} \\ 20\%: & \text{Gore} \succ \text{Bush} \succ \text{Nader} \\ 11\%: & \text{Nader} \succ \text{Gore} \succ \text{Bush} \end{array}$$

So even if nobody is cheating, Bush will win in a plurality contest.

Issue: In a *pairwise contest*, Gore would have defeated anyone.

Issue II: It would have been in the interest of the Nader supporters to *manipulate*, i.e. to misrepresent their preferences.

Condorcet Paradox

In 1785 the Marquis de Condorcet noticed a problem ...

$$\begin{array}{ll} \text{Agent 1:} & A \succ B \succ C \\ \text{Agent 2:} & B \succ C \succ A \\ \text{Agent 3:} & C \succ A \succ B \end{array}$$

How should we *aggregate the individual preferences* of these three agents into a *social preference* ordering?

A beats B and B beats C in pairwise contests. So probably we should have $A \succ B$ and $B \succ C$ in the social preference ordering. And by transitivity also $A \succ C$. But C beats A in a pairwise contest. This is known as the *Condorcet paradox*.

M. le Marquis de Condorcet. *Essai sur l'application de l'analyse à la probabilté des décisions rendues a la pluralité des voix.* Paris, 1785

Arrow's Impossibility Theorem

This is probably the most famous theorem in social choice theory. It was first proved by Kenneth J. Arrow in his 1951 PhD thesis. He later received the Nobel Prize in Economic Sciences in 1972.

The theorem shows that there can be no mechanism for aggregating individual preferences into a social preference that would simultaneously satisfy a small number of natural and seemingly innocent axioms.

Our exposition of the theorem is taken from Barberà (1980); the proof closely follows Geanakoplos (2005).

K.J. Arrow. *Social Choice and Individual Values*. 2nd edition, Wiley, 1963.

S. Barberà (1980). Pivotal Voters: A New Proof of Arrow's Theorem. *Economics Letters*, 6(1):13–16, 1980.

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

Setting

- Finite set of *alternatives* A.
- Finite set of *individuals* $I = \{1, \ldots, n\}$.
- A *preference ordering* is a strict linear order on A.
  The set of all such preference orderings is denoted $\mathcal{P}$.

  Each individual i has an *individual* preference ordering P_i, and we will try to find a *social* preference ordering P.
- A *preference profile* $\langle P_1, \ldots, P_n \rangle \in \mathcal{P}^n$ consists of a preference ordering for each individual.
- A *social welfare function* (SWF) is a mapping from preference profiles to social preference orderings: it specifies what preferences society should adopt for any given situation.

Axioms

It seems reasonable to postulate that any SWF should satisfy the following list of axioms:

- **(PAR)** The SWF should satisfy the *Pareto condition:* if every individual prefers x over y, then so should society.

$$(\forall \boldsymbol{P} \in \mathcal{P}^n)(\forall x, y \in A)[[(\forall i \in I) x P_i y] \rightarrow xPy]$$

- **(IIA)** The SWF should satisfy *independence of irrelevant alternatives:* social preference of x over y should not be affected if individuals change their preferences over other alternatives.

$$(\forall \boldsymbol{P}, \boldsymbol{P}' \in \mathcal{P}^n)(\forall x, y \in A)[[(\forall i \in I)(xP_i y \leftrightarrow xP_i' y)] \rightarrow (xPy \leftrightarrow xP'y)]$$

- **(ND)** The SWF should be *non-dictatorial:* no single individual should be able to impose a social preference ordering.

$$\neg(\exists i \in I)(\forall x, y \in A)(\forall \boldsymbol{P} \in \mathcal{P}^n)[xP_i y \rightarrow xPy]$$

The Result

Theorem 1 (Arrow, 1951) *If $|A| > 2$, then there exists no SWF that would simultaneously satisfy all of (PAR), (IIA) and (ND).*

Observe that if there are just two alternatives ($|A| = 2$), then it is easy to find an SWF that satisfies all three axioms (at least for an odd number of individuals): simply let the alternative preferred by the *majority* of individuals also be the socially preferred alternative.

Now for the proof . . .

Extremal Lemma

Assume (PAR) and (IIA) are satisfied. Let b be any alternative.

Claim: For any profile in which b is ranked either top or bottom by every individual, society must do the same.

Proof: Suppose otherwise; that is, suppose b is ranked either top or bottom by every individual, but not by society.

(1) Then aPb and bPc for distinct alternatives a, b, c and the social preference ordering P.

(2) By (IIA), this continues to hold if we move every c above a for every individual, as doing so does not affect the extremal b.

(3) By transitivity of P, we get aPc.

(4) But by (PAR), we get cPa. Contradiction. ✓

Existence of an Extremal Pivotal Individual

Fix some alternative b. We call an individual *extremal pivotal* iff it can move b from the bottom to the top of the social preference ordering (for some profile).

Claim: There exists an extremal pivotal individual.

Proof: Start with a profile where every individual puts b at the bottom. By (PAR), so does society.

Then let the individuals change their preferences one by one, moving b from the bottom to the top.

By the Extremal Lemma, there must be a point when the change in preference of a particular individual causes b to rise from the bottom to the top in the social ordering. ✓

Call the profile just before the switch in the social ordering occurred *Profile I*, and the one just after the switch *Profile II*.

Dictatorship: Case 1

Let i be the extremal pivotal individual (for alternative b). The existence of i is guaranteed by our previous argument.

Claim: Individual i can dictate the social ordering with respect to any alternatives a, c different from b.

Proof: Suppose i wants to place a above c.

Let *Profile III* be like *Profile II*, except that (1) i makes a its top choice (that is, aP_ibP_ic), and (2) all the others can rearrange their relative rankings of a and c as they please.

Observe that in *Profile III* all relative rankings for a, b are as in *Profile I*. So by (IIA), the social rankings must coincide: aPb.

Also observe that in *Profile III* all relative rankings for b, c are as in *Profile II*. So by (IIA), the social rankings must coincide: bPc.

By transitivity, we get aPc. By (IIA), this continues to hold if others change their relative ranking of alternatives other than a, c. ✓

Dictatorship: Case 2

Let b and i be defined as before.

Claim: Individual i can also dictate the social ordering with respect to b and any other alternative a.

Proof: We can use a similar construction as before to show that for a given alternative c, there must be an individual j that can dictate the relative social ordering of a and b (both different from c).

But at least in *Profiles I* and *II*, i can dictate the relative social ranking of a and b. As there can be at most one dictator in any situation, we get $i = j$. ✓

So individual i will be a *dictator* for *any* two alternatives.
This contradicts (ND), and Arrow's Theorem follows.

Monday Morning Quiz

A social welfare function satisfies *non-imposition* (NI) if any social preference ordering is achievable by *some* preference profile:

$$(\forall P \in \mathcal{P})(\exists \boldsymbol{P}' \in \mathcal{P}^n)(\forall x, y \in A)[xPy \leftrightarrow xP'y]$$

That is, a SWF satisfying (NI) doesn't impose any restrictions that would *a priori* exclude a particular social preference ordering.

Prove the following two claims:

- The Pareto condition (PAR) implies (NI).
- Arrow's Theorem breaks down if we replace (PAR) by (NI).

Recall Arrow's Theorem: for more than two alternatives, there exists not SWF that satisfies all of (PAR), (IIA) and (ND).

Computational Social Choice

Computational social choice studies collective decision making, with an emphasis on computational aspects. Work in COMSOC can be broadly classified along two dimensions —

The kind of social choice problem studied, e.g.:

- aggregating individual preferences into a collective ordering
- electing a winner given individual preferences over candidates
- fairly dividing a cake given individual tastes

The kind computational technique employed, e.g.:

- algorithm design to implement complex mechanisms
- complexity theory to understand limitations
- logical modelling to fully formalise intuitions

The next few slides are a collection of examples, followed by an overview of the rest of the course.

Electing a Committee

We have already seen that voting can be rather complicated: the election winner may be less popular than some other candidate; manipulation may be encouraged by the voting rule ... here is a further difficulty, this time of a computational nature.

Suppose we have to elect a *committee* (not just a single candidate):

- If there are k *seats* to be filled from a pool of m *candidates*, then there are $\binom{m}{k}$ possible outcomes.
- For $k = 5$ and $m = 12$, for instance, that's 792 alternatives.
- The domain of alternatives has a *combinatorial structure*.

It does not seem reasonable to ask voters to submit their full preferences over all alternatives to the collective decision making mechanism. What would be a reasonable form of balloting?

Mechanism Design

We have seen that *manipulation* is a serious problem in voting. In domains other than voting we can sometimes do better.

Suppose we want to sell a single item in an auction.

- *First-price sealed-bid auction:* each bidder submits an offer in a sealed envelope; highest bidder wins and pays what they offered
- *Vickrey auction:* each bidder submits an offer in a sealed envelope; highest bidder wins but pays *second highest price*

In the Vickrey auction each bidder has an incentive to submit their *truthful valuation* of the item!

W. Vickrey. Counterspeculation, Auctions, and Competitive Sealed Tenders. *Journal of Finance* 16(1):8–37, 1961.

Judgement Aggregation

Preferences are not the only structures that we may wish to aggregate. JA studies the aggregation of judgements on logically inter-connected propositions. Example:

	A	B	C
Judge 1:	yes	yes	yes
Judge 2:	no	yes	no
Judge 3:	yes	no	no
Majority:	yes	yes	no

A: witness is reliable

B: *if* witness is reliable *then* guilty

C: guilty

note that $A \wedge B \rightarrow C$

While each individual set of judgements is logically consistent, the collective judgement produced by the majority rule is not.

Ch. List and Ph. Pettit. Aggregating Sets of Judgments: Two Impossibility Results Compared. *Synthese* 140(1–2):207–235, 2004.

Earth Observation Satellites

Suppose our individual agents are representatives of different European countries that have jointly funded a new Earth Observation Satellite (EOS). Now they are requesting certain photos to be taken by the EOS, but due to physical constraints not all requests can be honoured . . .

Allocations should be both *efficient* and *fair*:

- The satellite should not be underexploited.
- Each agent should get a return on investment that is at least roughly proportional to its financial contribution.

M. Lemaître, G. Verfaillie, and N. Bataille. *Exploiting a Common Property Resource under a Fairness Constraint: A Case Study*. Proc. IJCAI-1999.

Efficiency and Fairness

When assessing the quality of an allocation (or any other decision) we can distinguish two types of indicators of *social welfare*.

Aspects of *efficiency* (*not* in the computational sense) include:

- The chosen agreement should be such that there is no alternative agreement that would be better for some and not worse for any of the other agents (*Pareto optimality*).
- If preferences are quantitative, the sum of all payoffs should be as high as possible (*utilitarianism*).

Aspects of *fairness* include:

- The agent that is going to be worst off should be as well off as possible (*egalitarianism*).
- No agent should prefer to take the bundle allocated to one of its peers rather than keeping their own (*envy-freeness*).

How do we formalise this? How do we compute optimal solutions?

Collective Decision Making

- *Preference aggregation:* aggregate a profile of individual preferences into a collective preference, i.e. a ranking of the alternatives form a social point of view.
  - we can also aggregate things other than preferences: *judgement aggregation*, *belief merging*, ...
- *Voting:* select a winner (or a group of winners) given a profile of individual preferences over candidates.
  - could be seen as a preference aggregation problem, where we only care about the top choice in the collective preference
- *Fair division and resource allocation:* select an allocation of goods to agents given their individual preferences.
  - could be seen as a voting problem, but that would mean ignoring useful information (e.g. an agent will typically be indifferent between allocations giving it the same bundle)

Logic and Computation

- Much work in SCT is axiomatic (e.g. Arrow's Theorem). While the field is mathematically *rigorous*, it is not (yet!) *formal*.
- Some research is about aggregating structures of a logical nature: *judgement aggregation*, *belief merging*.
- Tools from *knowledge representation* are useful for modelling preferences for social choice in combinatorial domains.
- In the long run it may become possible to specify social mechanisms in a suitable logic and to automatically verify certain properties ("*social software*").
- Some very involved social choice mechanisms require careful *algorithm design* to become usable.
- Tools from *complexity theory* can help understand limitations and design mechanisms that are hard to manipulate.

Tuesday

Tuesday will be an introduction to *voting theory*, with some examples for applications of *complexity theory* in that area.

- Many different voting procedures, such as the plurality rule, the Borda count, approval voting, single transferable vote, ...
- Discussion of properties ("axioms") we would like to see satisfied by a voting procedure
- May's Theorem and Gibbard-Satterthwaite Theorem
- Complexity as a barrier against manipulation: can we make it computationally intractable to manipulate a voting procedure?
- Other applications of complexity theory, e.g. analysis of the complexity of computing winners for complicated voting rules

Wednesday

Collective decision making can be particularly challenging when the alternatives to be decided upon have a combinatorial structure.

On Wednesday we will discuss *preference representation* languages for and *voting* in *combinatorial domains*.

- Overview of languages for representing both ordinal preference relations and utility functions (cardinal preference structures)
- Exemplification of properties such as expressivity, succinctness and complexity in the context of a specific language
- Approaches to voting in combinatorial domains: to what extent can we vote issue-by-issue?

Much of this takes inspiration from work coming out of the KRR (knowledge representation & reasoning) community.

Thursday

On Thursday we will start discussing *fair division*, focusing on *cake-cutting procedures* for allocating a single divisible good.

- Notions of fairness: proportionality and envy-freeness
- Overview of classical cake-cutting procedures from the literature, highlighting open problems
- Complexity of cake-cutting: how many cuts do we need to make a fair division?
- Pointers to literature on logical modelling of fair division and similar problems ("social software")

Friday

On Friday we will start with an introduction to *welfare economics*, briefly look into some related complexity questions, and then study *distributed mechanisms* for fairly allocating indivisible goods.

- Social welfare orderings and collective utility functions
- Complexity of finding a socially optimal allocation
- Design of interaction protocols that permit convergence to socially optimal allocations in distributed mechanisms

Literature

There are several textbooks on (classical) social choice theory in which you can find an exposition of Arrow's Theorem, for example:

- W. Gaertner. *A Primer in Social Choice Theory.* Oxford University Press, 2007.

For a tentative overview of the COMSOC research area, have a look at this survey paper:

- Y. Chevaleyre, U. Endriss, N. Maudet, and J. Lang. *A Short Introduction to Computational Social Choice.* SOFSEM-2007.