Computational Social Choice: Spring 2009

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

Introduction

The course will cover issues at the interface of *computer science* (including *logic*, *multiagent systems* and *artificial intelligence*) and *mathematical economics* (including *social choice theory*, *game theory* and *decision theory*).

There has been a recent *trend* towards research of this sort. The broad philosophy is generally the same, but people have been using different names to identify various flavours of this kind of work, e.g.:

- Algorithmic Game Theory
- Social Software
- *and:* Computational Social Choice

Very few specific *prerequisites* are required to follow the course. Nevertheless, we will frequently touch upon *current research* issues.

Organisational Matters

- **Lecturer:** Ulle Endriss (`u.endriss@uva.nl`), Room P.316
- **TA:** Umberto Grandi (`u.grandi@uva.nl`), Room P.311
- **Timetable:** Wednesdays 15–17 (+ two tutorials)
- **Examination:** There will be several *coursework assignments* on the material covered in the course. In the second block, every student will have to study a *recent paper*, write a short essay on the topic, and present their findings in a talk.
- **Website:** Lecture slides, coursework assignments, and other important information will be posted on the course website:

  `http://www.illc.uva.nl/~ulle/teaching/comsoc/`
- **Seminars:** There are occasional talks at the ILLC that are directly relevant to the course and that you are welcome to attend (e.g., at the Computational Social Choice Seminar).

Prerequisites

There are no formal prerequisites. But: you should be comfortable with *formal* material and you will be asked to *prove* stuff.

There are two areas for which we will assume some background knowledge that some of you may not yet have. This material will be covered in *tutorials* in the first two weeks:

- **Game Theory:** non-cooperative games in strategic form; Pareto optimal outcomes; dominant strategies; pure and mixed Nash equilibria; computing Nash equilibria for small games
- **Complexity Theory:** definition of complexity classes such as P and NP; completeness with respect to a complexity class; proving NP-completeness via reduction

Related Courses

- Cooperative Games
  Krzysztof Apt
- Games and Complexity
  Peter van Emde Boas
- Introduction to Game Theory (in autumn)
  Peter van Emde Boas
- Logic and Games (or similar, sometimes offered in autumn)
  Johan van Benthem
- Multiagent Systems and Distributed AI (MSc AI)
  Shimon Whiteson

Plan for Today

- **Part I:** Introduction to the main topics of the course
- **Part II:** Arrow's Impossibility Theorem
  (as an example for a classical result in social choice theory)

Part I: Course Topics

Collective Decision Making

This course is about *collective decision making:* How can we map the individual preferences of a group of agents into a joint decision?

Next we will see some examples, problems, ideas, paradoxes, or just *issues* that illustrate the main question addressed in the course:

▶ *How does collective decision making work?*

The remainder of the course will then be devoted to developing these rather vague ideas in a rigorous manner.

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:

$$49\%: \text{Bush} \succ \text{Gore} \succ \text{Nader}$$
$$20\%: \text{Gore} \succ \text{Nader} \succ \text{Bush}$$
$$20\%: \text{Gore} \succ \text{Bush} \succ \text{Nader}$$
$$11\%: \text{Nader} \succ \text{Gore} \succ \text{Bush}$$

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 ...

Agent 1: $A \succ B \succ C$
Agent 2: $B \succ C \succ A$
Agent 3: $C \succ A \succ B$

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

A majority prefers A over B and a majority also prefers B over C, but then again a majority prefers C over A.

So the social preference ordering induced by the seemingly natural majority rule fails to be rational (it's not transitive).

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

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.

Vickrey Auctions

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!

William Vickrey received the 1996 Nobel Prize in Economic Sciences for "contributions to the economic theory of incentives".

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

Voting and Complexity

In a sense to be made precise (next week), the positive result for the Vickrey auction cannot be transferred to the domain of voting: it is impossible to incentivice agents to always vote truthfully.

But we can try to use *complexity* as a *barrier against manipulation:*

- For some voting rules it is *computationally intractable* to compute my *best* (insincere) vote (even if I correctly guess everybody else's vote)—so I should probably vote sincerely.

Other applications of complexity theory in voting:

- People have come up with pretty complicated voting rules. What is the complexity of determining the *winner*?
- How hard is it to check whether a given candidate can *possibly* still win after part of the ballots have been counted?
- What is the complexity of *bribery*? of *control* by the chair?

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?

Preference Representation Languages

We will look into several languages for representing preferences. When choosing a language, we should consider these criteria:

- *Cognitive relevance:* How close is a given language to the way in which humans would express their preferences?
- *Elicitation:* How difficult is it to elicit the preferences of an agent so as to represent them in the chosen language?
- *Expressive power:* Can the chosen language encode all the preference structures we are interested in?
- *Succinctness:* How compact is the representation of (typical) preferences? Is one language more succinct than another?
- *Complexity:* What is the computational complexity of related decision problems, such as comparing two alternatives?

Earth Observation Satellites

Our agents are representatives of different European countries that have jointly funded a new Earth Observation Satellite (EOS). Now the agents 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 their financial contribution.

This is an example for a *multiagent resource allocation* problem.

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 their peers rather than keeping their own (*envy-freeness*).

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

Summary

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 of computational technique employed, e.g.:

- algorithm design to implement complex mechanisms
- complexity theory to understand limitations
- logical modelling to fully formalise intuitions
- KRR techniques to efficiently model social choice problems

Part II: Arrow's Theorem

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 by $\mathcal{P}$.
  Each individual i has an *individual* preference ordering P_i, and we will try to find a *social* preference ordering P.
- A *preference profile* $\boldsymbol{P} = \langle P_1, \ldots, P_n \rangle \in \mathcal{P}^n$ 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.
- Remark: We implicitly assume that *any* individual preference orderings are possible (*universal domain* assumption).

Axioms

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

- **(WP)** The SWF should satisfy the *weak Pareto condition* (aka. *unanimity*): if everyone prefers x over y, then so should society.

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

- **(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.

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

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

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

The Result

Theorem 1 (Arrow, 1951) *For three or more alternatives, there exists no SWF that satisfies all of (WP), (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 (WP) 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, applied to (1), we get aPc.

(4) But by (WP), applied to (2), we get cPa. Contradiction. ✓

Existence of an Extremal Pivotal Individual

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

Claim: There exists an extremal pivotal individual i.

Proof: Start with a profile where every individual puts b at the bottom. By (WP), 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 and (WP), 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 have rearranged 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.

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.
- A.D. Taylor. *Social Choice and the Mathematics of Manipulation*. Cambridge University Press, 2005.

(This is optional reading.)

Literature

There is no textbook or similar for COMSOC. I will recommend specific papers or book chapters in each lecture. For the general feeling, you should also browse through some of these:

- Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. *A Short Introduction to Computational Social Choice*. Proc. SOFSEM-2007.
- 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.
- Y. Chevaleyre *et al*. Issues in Multiagent Resource Allocation. *Informatica*, 30:3–31, 2006.
- C.H. Papadimitriou. *Algorithms, Games, and the Internet*. Proc. STOC-2001. (about algorithmic game theory)

What next?

The main topics that we are going to cover in this course are:

- (Computational Issues in) Voting Theory
- Preference Handling in Combinatorial Domains
- Multiagent Resource Allocation and Fair Division
- Combinatorial Auctions and Mechanism Design