Voting on Actions with Uncertain Outcomes

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

Outline

General Introduction to Computational Social Choice

- What is Social Choice Theory?
- How does it relate to Computer Science / AI / Logic?

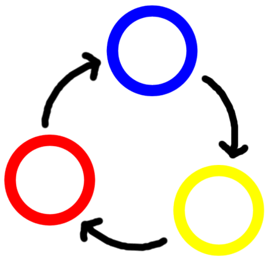
Main Topic: Voting on Actions with Uncertain Outcomes

- New Model
- Examples
- Initial Results

Social Choice and the Condorcet Paradox

Social Choice Theory asks: how should we aggregate the preferences of the members of a group to obtain a "social preference"?

Expert 1: 🔵 $\succ$ 🟡 $\succ$ 🔴
Expert 2: 🟡 $\succ$ 🔴 $\succ$ 🔵
Expert 3: 🔴 $\succ$ 🔵 $\succ$ 🟡
Expert 4: 🔴 $\succ$ 🔵 $\succ$ 🟡
Expert 5: 🟡 $\succ$ 🔴 $\succ$ 🔵



Marie Jean Antoine Nicolas de Caritat (1743–1794), better known as the **Marquis de Condorcet**: Highly influential Mathematician, Philosopher, Political Scientist, Political Activist. Observed that the *majority rule* may produce inconsistent outcomes ("Condorcet Paradox").

A Classic: Arrow's Impossibility Theorem

In 1951, K.J. Arrow published his famous *Impossibility Theorem:*

Any preference aggregation mechanism for *three* or more alternatives that satisfies the axioms of *unanimity* and *IIA* must be *dictatorial*.

- Unanimity: if everyone says $A \succ B$, then so should society.
- Independence of Irrelevant Alternatives (IIA): if society says $A \succ B$ and someone changes their ranking of C, then society should still say $A \succ B$.

Kenneth J. Arrow (born 1921): American Economist; Professor Emeritus of Economics at Stanford; Nobel Prize in Economics 1972 (youngest recipient ever). His 1951 PhD thesis started modern Social Choice Theory. Google Scholar lists 13,167 citations of the thesis.

Modern Applications of Social Choice Theory

Social choice theory has natural applications in computer science:

- *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
- *Multiagent Systems:* to aggregate the beliefs + to coordinate the actions of groups of autonomous software agents

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

CS/AI/Logic for Social Choice Theory

Vice versa, techniques from computer science are useful for advancing the state of the art in social choice theory:

- *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

F. Brandt, V. Conitzer, and U. Endriss. Computational Social Choice. In G. Weiss (ed.), *Multiagent Systems*, pages 213–283. MIT Press, 2013.

Voting on Actions with Uncertain Outcomes

Scenario: A group of *agents* have to decide on an *action* to take, but they are *uncertain* about the *effects* of the available actions. Each agent has *preferences* over possible outcomes (i.e., over effects of actions, *not* over actions themselves) and each of them has *beliefs* regarding the likely effects of actions. We need to *aggregate both* of these forms of information to come to a socially desirable solution.

- ▶ What *method* should we use?
  But first: How should we *model* this?

I do *not* want to model it in terms of *expected utility* etc.:

- Agents might not be able to assign precise *utilities* to outcomes
- Agents might not be able to assign precise *probabilities* to events

Instead, I want a simple qualitative model.

The Model

The world:

- Deterministic finite state machine: *states* and *actions*, as well as a *transition function* mapping any state/action pair to a next state

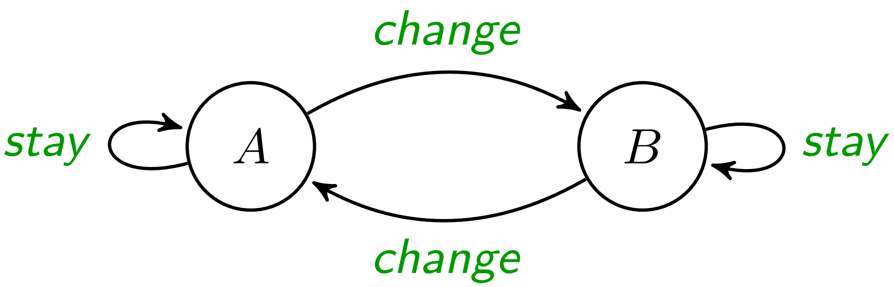
This description of the world is known to all agents (no uncertainty).

Each of a finite set of *agents* has her own

- *Beliefs:* modelled as a *subset* of states she considers plausible current states (*before* execution of the action)
- *Preferences:* modelled as a *linear order* over the set of states (*after* execution of the action)

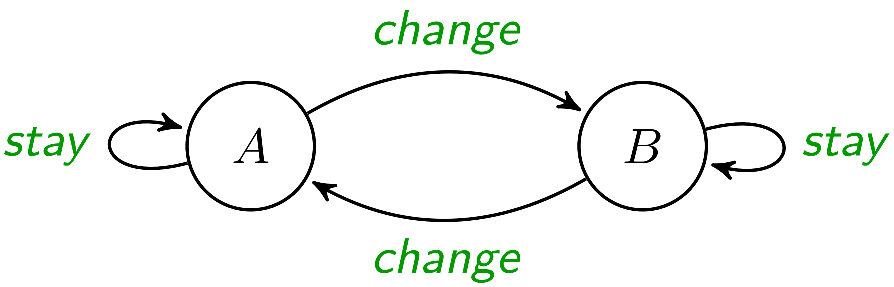
Discussion: uncertain about effect of action *vs.* uncertain about current state

Example



	Belief	Preference	Action
Agent 1	A	$A \succ B$	*stay*
Agent 2	A	$B \succ A$	*change*
Agent 3	B	$B \succ A$	*stay*
Collective			*stay*

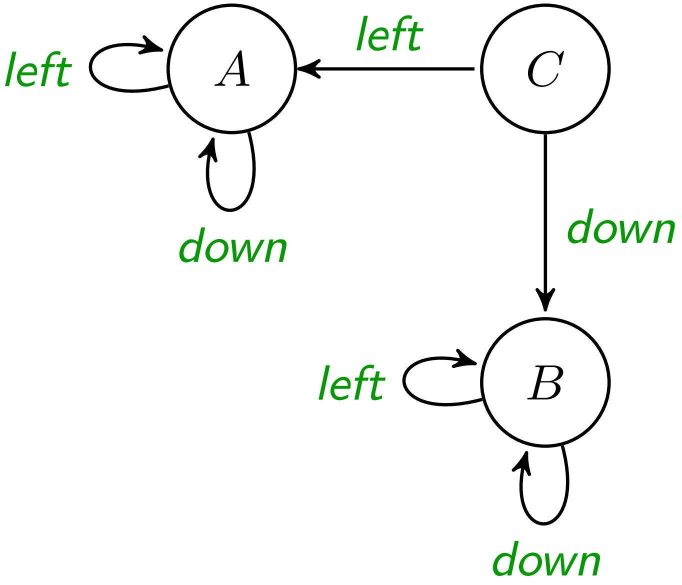
The Paradox of Individual Uncertainty Resolution



	Belief	Preference	Action
Agent 1	A	$A \succ B$	*stay*
Agent 2	A	$B \succ A$	*change*
Agent 3	B	$B \succ A$	*stay*
Collective			*stay*

	Belief	Preference	Action
Agent 1	A	$A \succ B$	
Agent 2	A	$B \succ A$	
Agent 3	B	$B \succ A$	
Collective	A	$B \succ A$	*change*

	Belief	Preference	Action
Agents 1–9	A or C	$A \succ C \succ B$	
Agent 10	A or B	$B \succ C \succ A$	
Collective	A	$A \succ C \succ B$	*down*

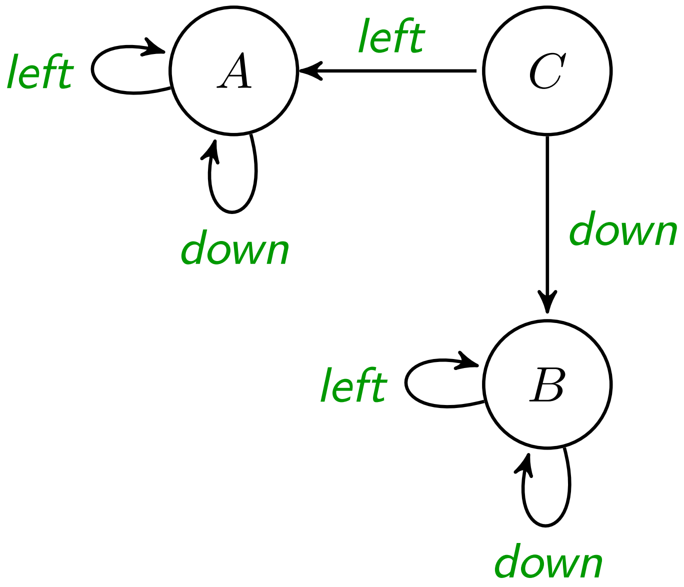


[break ties in favour of *down*]

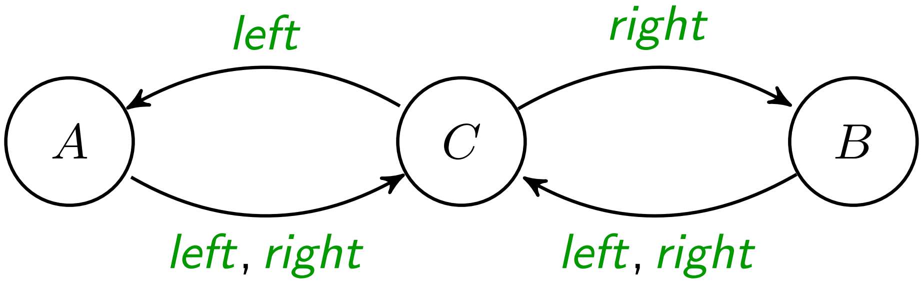
The Paradox of Early Collective Uncertainty Resolution

	Belief	Preference	Action
Agents 1–9	A or C	$A \succ C \succ B$	
Agent 10	A or B	$B \succ C \succ A$	
Collective	A	$A \succ C \succ B$	*down*

	Belief	Preference	Action
Agents 1–9	A or C	$A \succ C \succ B$	
Agent 10	A or B	$B \succ C \succ A$	
Collective	A [or C]	$A \succ C \succ B$	*left*



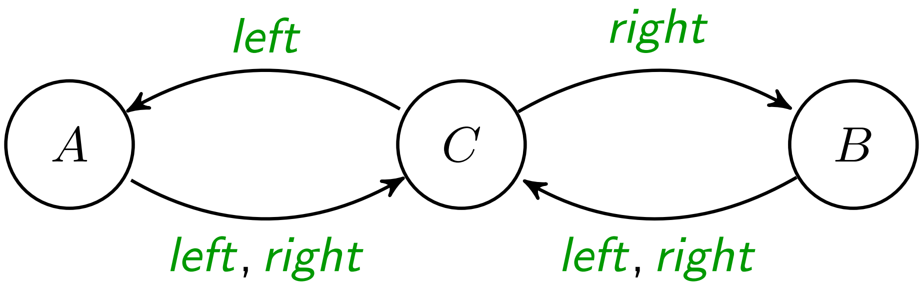
[break ties in favour of *down*]



	Belief	Preference	Action
Agents 1–2	A or C	$A \succ C \succ B$	
Agents 3–5	B or C	$B \succ A \succ C$	
Collective	C	$A \succ B \succ C$	*left*

[aggregate preferences using *Borda*]

The Paradox of Late Collective Uncertainty Resolution



	Belief	Preference	Action
Agents 1–2	A or C	$A \succ C \succ B$	
Agents 3–5	B or C	$B \succ A \succ C$	
Collective	C	$A \succ B \succ C$	*left*

	Belief	Preference	Action
Agents 1–2	A or C	$A \succ \cancel{C} \succ B$	
Agents 3–5	B or C	$B \succ A \succ \cancel{C}$	
Collective	C	$B \succ A$	*right*

[aggregate preferences using *Borda*]

Preference Aggregation in Isolation

Disregard the belief component for the moment.

How to *aggregate* the individual *preferences* into a collective order?

This is the classical problem of social choice theory:

- no perfect solution (see, e.g., Arrow's Theorem)
- but trade-offs to be made are well-understood
- many social choice theorists would recommend the *Kemeny* rule

Belief Aggregation in Isolation

Now disregard the preference component.

Recall: individual *beliefs* are modelled as *sets of plausible states*.

So a *belief aggregator* will be a function mapping any profile of sets of states into a single (collective) set of states.

This does not correspond to any standard problem in SCT.
What's best depends on our interpretation of the sets supplied:

- If agents report *knowledge*, then all individual belief sets must include the true state ⇒ take a *subset of their intersection*.

  Small characterisation result: if you want *neutrality*, then you must choose exactly the *intersection* (no proper subset).

- If agents merely report *beliefs*, then interesting aggregators include *approval voting* and the *mean-based rule*.

Integration of the Two Aggregation Outcomes

For our original problem of voting under uncertainty, one approach is:

(1) Use your favourite method of *preference aggregation* to obtain a single (collective) *preference order* over outcomes.

(2) Use your favourite method of *belief aggregation* to obtain a single (collective) *belief set* regarding plausible current states.

(3) Now *combine the two* to pick the best action.

That is: at this point, treat it as a *single-agent problem*.

Note: This is not the only possible approach.

Desiderata for the Single-Agent Case

Given a *set of plausible states* and a *preference order* on outcomes, how should you *rank* the available *actions*?

Two ways of approaching this: consider the set of possible outcomes as a whole, or consider possible states case by case.

- *Outcome Dominance Axiom:* Every given action induces a set of *plausible outcomes*. Prefer action α over β if you'd rather have someone pick from the set induced by α than the set induced by β.

  $$\delta(Q,\alpha) \text{ \textit{Gärdenfors-dominates} } \delta(Q,\beta) \;\Rightarrow\; \alpha \succ_Q \beta$$

- *Casewise Dominance Axiom:* Prefer action α over β if α gives at least as good* a result as β for every state considered plausible.

  $$\delta(q,\alpha) \succcurlyeq \delta(q,\beta) \text{ for all } q \in Q \text{ [}^\star\text{strictly for some]} \;\Rightarrow\; \alpha \succ_Q \beta$$

Can we find an *action ranking function* that satisfies these axioms?

An Impossibility Theorem

Much weaker than our outcome dominance axiom:

- *Outcome Relevance Axiom:* remain indifferent between actions α and β if they give rise to the same set of possible outcomes.

$$\delta(Q,\alpha) = \delta(Q,\beta) \;\Rightarrow\; \alpha \sim_Q \beta$$

Still, bad news:

> *There exists no action ranking function that satisfies both casewise dominance and outcome relevance.*

Recall: *casewise dominance* means that we prefer α over β if α gives at least as good* a result as β for every state considered plausible.

Last Slide

After a few general words on *computational social choice* I have

- introduced a simple model for *voting under uncertainty*,
- demonstrated its interestingness through *three paradoxes*,
- briefly discussed possible *aggregation methods*, and
- presented an *impossibility result* for the single agent case.

Outlook: The seemingly weak *outcome relevance axiom* actually is much *too strong*. So not all hope is lost. But devising good methods of aggregation is still a serious challenge.

U. Endriss. Voting on Actions with Uncertain Outcomes. *Proc. 3rd International Conference on Algorithmic Decision Theory (ADT-2013)*. Springer, 2013.