Voting on Actions with Uncertain Outcomes

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

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

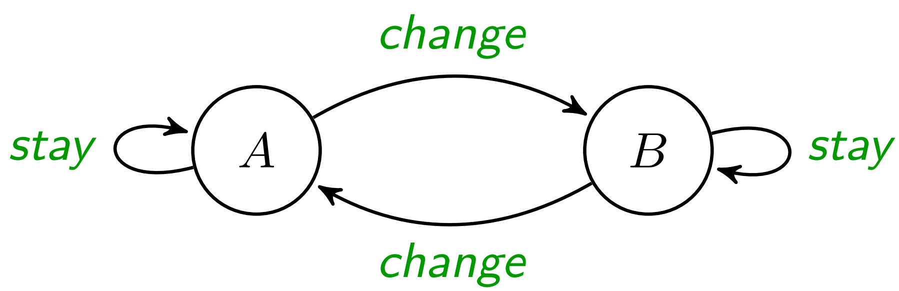
This description of the world is known to all agents (no uncertainty).

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)

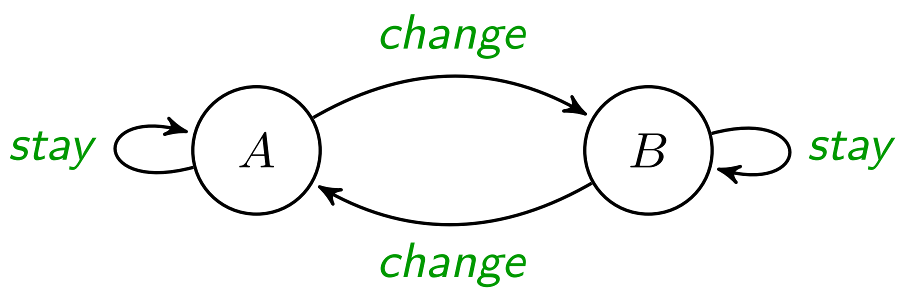
Discussion: uncertain about effect of action *vs.* uncertain about current state

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*

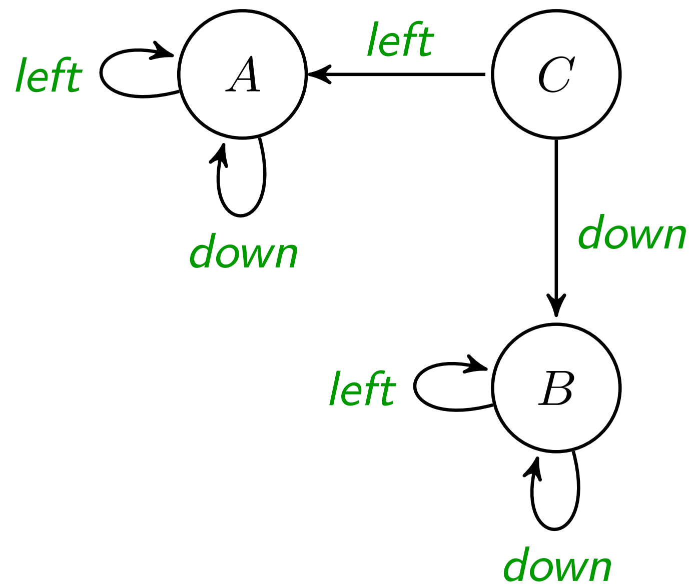
The Paradox of Individual Uncertainty Resolution



	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*

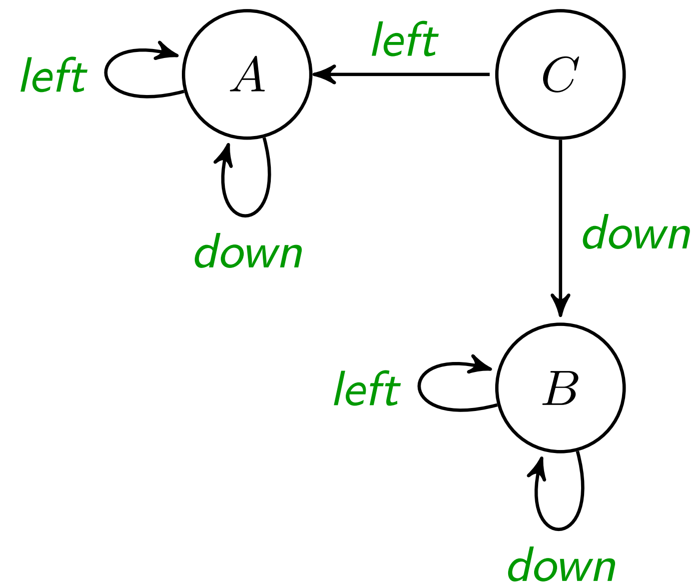


[break ties in favour of *down*]

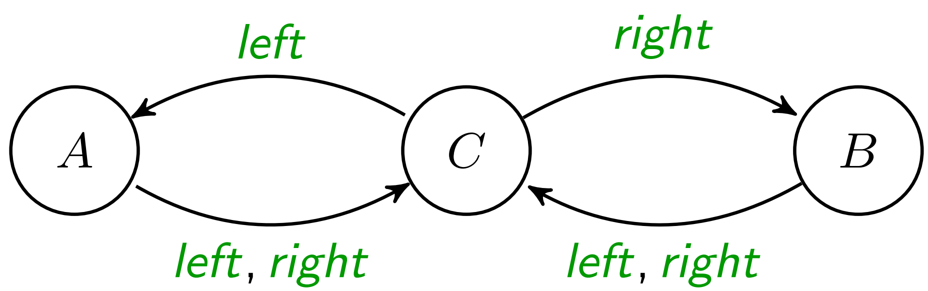
The Paradox of Early Collective Uncertainty Resolution

	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*



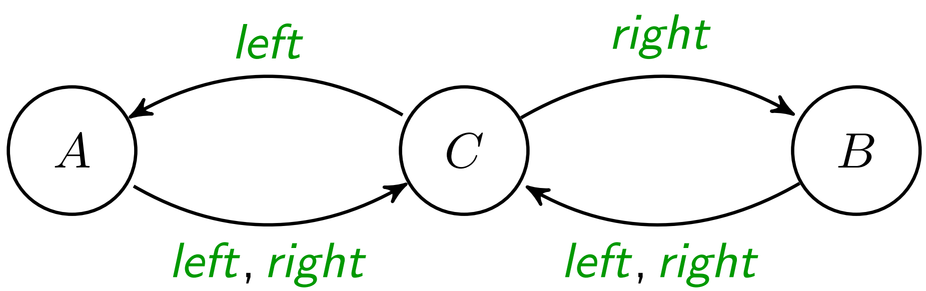
[break ties in favour of *down*]



	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	$A \succ$ ~~C~~ $\succ B$	
Agents 3–5	B or C	$B \succ A \succ$ ~~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 (but, e.g., Kemeny rule not too bad).

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.

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

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

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

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

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.