# Epistemic logics for time and space bounded reasoning

**Natasha Alechina**

University of Nottingham

joint work with
Brian Logan and Mark Jago (University of Nottingham),
Thomas Ågotnes (University of Bergen),
Piergiorgio Bertoli, Chiara Ghidini and Luciano Serafini (ITC-irst, Trento)

# Outline

- Epistemic logic and logical omniscience
- General idea (state transitions = inference steps)
- Example systems
- Model-checking
- Extended language for bounded memory: ‘the agent has at most n beliefs’

# Epistemic Logic

Logic describing knowledge or belief. Let $B$ be a belief operator; '$B\alpha'$ means 'the agent believes formula $\alpha$'.

Usually interpreted over possible worlds structures $M = (W, R, V)$, where $W$ is a non-empty set of possible worlds, $R \subseteq W^2$, and $V : W \times Prop \to \{\top, \bot\}$ a valuation of propositional variables.

- $M, w \models p$ iff $V(w, p) = \top$
- $M, w \models B\phi$ iff for all $v$ such that $R(w, v)$, $M, v \models \phi$.

Intended interpretation: $B\alpha$ is true in $w$ if in all worlds which are considered possible in $w$, $\alpha$ is true.

# Epistemic Logic: axioms

**K** $B(\phi \to \psi) \to (B\phi \to B\psi)$

**N** $\vdash \phi \Rightarrow \vdash B\phi$

Plus, sometimes add ability to introspect, consistency, true belief = knowledge (correspond to simple conditions on $R$):

**4** $B\phi \to BB\phi$

**5** $\neg B\phi \to B\neg B\phi$

**D** $\neg B\bot$

**T** $B\phi \to \phi$

# Logical omniscience

The agent believes all tautologies, and all classical consequences of its beliefs.

Is this a problem?

Depends what the logic is meant to be used for...

- to describe idealised rational agents (not a problem)
- to describe non-idealised agents which don't reason; their actions depend on very simple atomic beliefs, and don't depend on belief or otherwise in tautologies (not a problem)
- to describe agents which answer complicated queries using for example classical logic, or some other set of inference rules. Not a problem if every query is going to be computed in the end, and we are not worried about time.

# When is logical omniscience a problem

If we want to reason about knowledge/belief of agents which

- do reasoning
- have limited memory (e.g., can not possibly believe all tautologies even if they know the inference rules to derive them)
- act/reason in situations where time matters (the environment changes, or answers are required in limited time, etc.)

# In that case...

- Belief operator has to be defined syntactically (beliefs are formulas, not sets of possible words)
- Time and memory have to be accounted for in the logic.

The purpose of the talk is to show that, contrary to the widely held opinion, syntactic logics of belief can be interesting.

# General idea and similar approaches

Let us say that an agent has a finite (or even bounded size) set of beliefs, and can make transitions to new states by using its inference rules.

Konolige: agents parametrised by sets of inference rules, but their theories are closed with respect to those rules.

Perlis et al. (step logic). Agents have finite sets of beliefs; application of inference rules to the set of beliefs advances the clock.

Fagin, Halpern, Moses, Vardi, later Pucella and Halpern (deductive algorithmic knowledge): explicit beliefs are closed with respect to agent's deductive capabilities.

Ågotnes and Walicki: explicit knowledge of inference rules; ATEL-style formalism where 'action' of an agent is an application of an inference rule.

# Structures (single agent case)

The reasoner/agent has internal language is $L$ and a set of inference rules $\mathcal{R}$.

A structure $M = (S, R)$ consists of a non-empty set $S$ of states and a binary relation $R \subseteq S \times S$.

- Elements of $S$ are finite sets of $L$ formulas (beliefs).
- $R$ holds between $s$ and $s'$ if there exists a formula $\alpha \in L$ which is derivable from the formulas in $s$ by one application of rule n $\mathcal{R}$, and $s' = s \cup \{\alpha\}$.

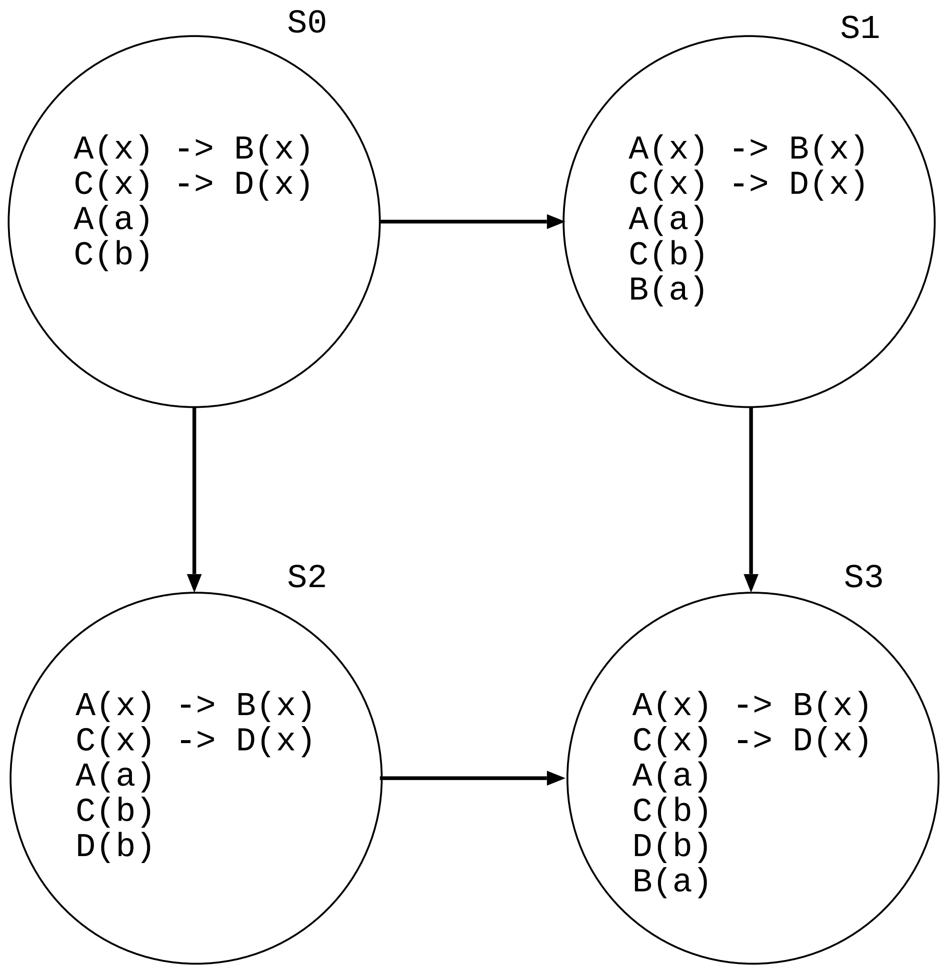
# Structures (example)

Suppose for example that $L$ is Horn clause rules and ground literals in some finite signature without functional symbols, and the agent has a single inference rule:

$$\frac{\forall \bar{x}(A_1, \ldots, A_n \rightarrow A), \quad \delta(A_1), \ldots, \delta(A_n)}{\delta(A)}$$

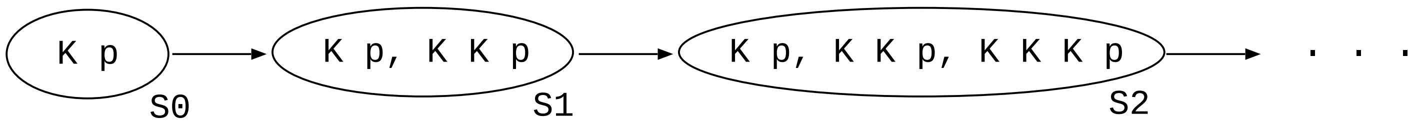
where $\delta$ is a substitution function.

Then $R(s, s')$ iff $s' = s \cup \delta(A)$ for some $A_1, \ldots, A_n \rightarrow A$, $\delta(A_1), \ldots, \delta(A_n) \in s$.


S0
A(x) -> B(x)
C(x) -> D(x)
A(a)
C(b)
S1
A(x) -> B(x)
C(x) -> D(x)
A(a)
C(b)
B(a)
S2
A(x) -> B(x)
C(x) -> D(x)
A(a)
C(b)
D(b)
S3
A(x) -> B(x)
C(x) -> D(x)
A(a)
C(b)
D(b)
B(a)

# Another example

Agent's language: epistemic propositional logic; agent's rules: introspection (from $K\phi$ conclude $KK\phi$).



# Language

The language of the epistemic logic $ML$ is defined relative to the agent's internal language $L$.

- if $\alpha$ is an $L$-formula, then $B\alpha$ ('the agent believes that $\alpha$') is an $ML(L)$ formula;
- if $\phi$ is an $ML(L)$ formula, then so are $\neg\phi$ and $\Diamond\phi$ ('in some successor state, $\phi$') and $\Box\phi$ ('in all successor states, $\phi$').
- if $\phi_1$ and $\phi_2$ are $ML(L)$ formulas, then so are $\phi_1 \wedge \phi_2$, $\phi_1 \rightarrow \phi_2$ etc.

# Semantics

$M, s \models \phi$ ($\phi$ is satisfied by $s \in S$ in $M = (S, R)$):

- $M, s \models B\alpha$ iff $\alpha \in s$
- $M, s \models \neg\phi$ iff $M, s \not\models \phi$
- $M, s \models \phi \wedge \psi$ iff $M, s \models \phi$ and $M, s \models \psi$
- $M, s \models \Diamond\phi$ iff there exists a $t \in S$ such that $R(s, t)$ and $M, t \models \phi$
- $M, s \models \Box\phi$ iff for all $t \in S$ such that $R(s, t)$, $M, t \models \phi$

# Axioms for $\Box$

Obviously, whatever conditions define the accessibility relation, we have the modal **K** axiom:

$$\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$$

and the necessitation rule:

$$\frac{\phi}{\Box\phi}$$

# Axioms connecting $\Box$ and $B$

If the reasoner does not discard beliefs, but only acquires new ones:

$$B\alpha \rightarrow \Box B\alpha$$

# Axioms connecting $\Box$ and $B$

At most one new belief is added in each transition:

$$\Diamond(B\alpha_1 \wedge B\alpha_2) \rightarrow (B\alpha_1 \vee B\alpha_2)$$

# Axioms connecting $\Box$ and $B$

The rest depends on the agent's inference rules; for example, to say that an agent can add new beliefs by

$$\frac{\alpha_1, \ldots, \alpha_n \rightarrow \alpha, \quad \alpha_1, \ldots, \alpha_n}{\alpha}$$

we can write

$$B(\alpha_1, \ldots, \alpha_n \rightarrow \alpha) \wedge B\alpha_1 \ldots \wedge B\alpha_n \rightarrow \Diamond B\alpha$$

# Axioms connecting $\Box$ and $B$

To say that an agent can add new beliefs by

$$\frac{\alpha_1, \quad \alpha_2}{\alpha_1 \wedge \alpha_2}$$

we can write

$$B\alpha_1 \wedge B\alpha_2 \rightarrow \Diamond B(\alpha_1 \wedge \alpha_2)$$

# Axioms connecting $\Box$ and $B$

But how do we say that *only* such transitions are possible?

It is easy in case of

$$\frac{\alpha_1, \quad \alpha_2}{\alpha_1 \wedge \alpha_2}$$

namely,

$$\Diamond B(\alpha_1 \wedge \alpha_2) \rightarrow B(\alpha_1 \wedge \alpha_2) \vee (B\alpha_1 \wedge B\alpha_2)$$

# Axioms connecting $\Box$ and $B$

However, to say that the only transitions which are possible are made by for example modus ponens

$$\frac{\alpha_1, \quad \alpha_1 \rightarrow \alpha_2}{\alpha_2}$$

we need an infinite disjunctions over all $\alpha_1$:

$$\Diamond B\alpha_2 \rightarrow B\alpha_2 \vee \bigvee_{\alpha_1 \in L} (B\alpha_1 \wedge B(\alpha_1 \rightarrow \alpha_2))$$

## From the modal logic point of view...

- 'Belief formulas' $B\alpha$ are essentially like propositional variables
- Conditions on accessibility relation $R$ are defined in terms of assignment to propositional variables
- In resulting modal logic, each state may satisfy only finitely many propositional variables; modal equivalence and bisimulation coincide (since assignment to variables determines which states are accessible)
- Depending on the agent's rules, the logic may or may not have the finite model property (e.g., 'modus ponens' agent has FMP and 'conjunction introduction' does not).

# Completeness and decidability results

- Complete axiomatisation of ‘rule-based agents’ with a fixed program (set of rules), see next slide. Joint work with Brian Logan and Mark Jago. Mark’s PhD thesis has more examples, also for communicating agents.
- Various examples of inference rules (MP, conjunction introduction, various communication rules): joint work with Thomas Ågotnes, Technical Report 304, Department of Informatics, University of Bergen.

# Logic of rule-based agents

We assume a fixed program $\mathcal{R}$ (and a finite domain of individuals, so the set of possible ground literals is also finite).

Condition on models:

- all states contains the same rules $\mathcal{R}$
- $R(s, s')$ iff for some rule $\lambda_1, \ldots, \lambda_n \rightarrow \lambda$ in $\mathcal{R}$ and some substitution $\delta$ such that $\delta(\lambda_1), \ldots, \delta(\lambda_n) \in s$, $s' = s \cup \delta(\lambda)$.

Note that it is possible that $s = s'$.

# Axiomatisation

**A1** $B\rho$, where $\rho \in \mathcal{R}$ (agent believes its rules)

**A2** $\neg B\rho$, where $\rho$ is an implication and $\rho \notin \mathcal{R}$ (agent only believes its rules)

**A3** $B\alpha \rightarrow \Box B\alpha$ (agents are monotonic reasoners)

**A4** $B(\lambda_1, \ldots, \lambda_n \rightarrow \lambda) \wedge B\delta(\lambda_1) \wedge \ldots \wedge B\delta(\lambda_n) \rightarrow \Diamond B\delta(\lambda)$ (if a rule matches, its consequent is added to some successor state)

**A5** $\Diamond(B\alpha \wedge B\beta) \rightarrow (B\alpha \vee B\beta)$ (at most one new belief is added in each transition)

**A6** $\Diamond B\alpha \rightarrow (B\alpha \;\vee \bigvee_{\lambda_1,\ldots,\lambda_n \rightarrow \lambda \in \mathcal{R}, \delta(\lambda)=\alpha} B\delta(\lambda_1) \wedge \ldots \wedge B\delta(\lambda_n))$ (new beliefs only arise as a result of firing a rule)

# Bounded memory

Suppose we want to model reasoners whose memory is not just finite, but of bounded size: contains at most $n$ beliefs.

This can be expressed by the following axiom schema:

$$B\alpha_1 \wedge \ldots \wedge B\alpha_n \rightarrow \neg B\alpha_{n+1}$$

(for distinct $\alpha_1, \ldots, \alpha_{n+1}$.)

# Bounded memory: non-monotonicity

If the agent has non-trivial inference rules, for example

$$B\alpha_1 \wedge B\alpha_2 \rightarrow \Diamond B(\alpha_1 \wedge \alpha_2)$$

then the bounded memory condition

$$B\alpha_1 \wedge \ldots \wedge B\alpha_n \rightarrow \neg B\alpha_{n+1}$$

is incompatible with monotonicity:

$$B\alpha \rightarrow \Box B\alpha$$

Either the reasoner becomes non-monotonic and may overwrite some belief with a newly derived belief in the next state; or inference steps are only possible when the memory is not full.

However, we cannot express that the memory is full in $ML(L)$ without an infinite disjunction.

# Bounded memory: overwriting formulas

We can express the property 'at most one formula becomes overwritten in each transition':

$$\Diamond(\neg B\alpha_1 \wedge \neg B\alpha_2) \rightarrow (\neg B\alpha_1 \vee \neg B\alpha_2)$$

But we cannot say (without an infinite disjunction): only if the memory is full (contains $n$ formulas), is one of the formulas overwritten.

# Model-checking space and time requirements

joint work with Piergiorgio Bertoli, Chiara Ghidini and Luciano Serafini from ITC-IRST in Trento.

Using MBP (Model-Based Planner, `http://sra.itc.it/tools/mbp/`) to establish how much memory is required for a certain derivation, and trade-off between memory and time (the number of steps in the derivation).

# Example problem

The agent can:

- read formulas one at a time from the knowledge base
- apply conjunction introduction
- apply modus ponens

The knowledge base contains $A, A \to B, A \wedge B \to C, B \wedge C \to D$; can the agent derive $D$ with memory of size 2?

# Idea of encoding

**States** are identified with assignments to the set of formulas $\Omega$ which are all subformulas of the knowledge base $KB$ and the goal formula $\phi_G$.

**Transition relation** is defined, for convenience, as having two arguments: the formula to be derived and the formula to be overwritten. For example, the precondition of $\mathbf{Read}(A, B)$ is that $A$ is in $KB$, and the postcondition is that in the next state, $A$ is true and $B$ is false (the agent believes $A$ and does not believe $B$). The precondition of $\mathbf{AND}(A_1 \wedge A_2, B)$ is that $A_1 \wedge A_2$ is in $\Omega$, and $A_1$ and $A_2$ are true.

# Success condition

In all the cases considered above (formulas in the next state are always consequences of the formulas in the previous state), the success condition is simple: The agent with beliefs $s_0$ can derive $\phi_G$, if from $s_0$ there is a path to a state containing $\phi_G$.

Things get more interesting when the agent does reasoning by cases; for example if it believes $A \vee B$ in $s$, then it has to explore states reachable from $s \cup \{A\}$ and states reachable from $s \cup \{B\}$. If a state containing $C$ is reachable, it does not mean that $C$ follows from $A \vee B$; it may be only reachable along the $A$-branch.

In this case, the agent with beliefs $s_0$ can derive $\phi_G$, if there is a plan (sequence of rules, including non-deterministic choice of a disjunct), all executions of which lead to a state containing $\phi_G$.

# Experiments

So far we only verified very small examples; see ITC-IRST Technical Report T05-10-03.

Future work: see how the approach scales, and improve memory model for non-deterministic reasoners (at the moment, they don't pay for remembering choice points).

# Extending the language

Since we cannot say 'the memory is full'...

In his PhD thesis, Thomas Ågotnes considered an epistemic operator $\bigtriangledown\{\alpha_1, \ldots, \alpha_n\}$ which means: the agent believes at most the formulas $\alpha_1, \ldots, \alpha_n$.

$M, s \models \bigtriangledown\{\alpha_1, \ldots, \alpha_n\}$ iff $s \subseteq \{\alpha_1, \ldots, \alpha_n\}$.

This is not definable using $B$.

Recently, we decided to study $min(n)$, for each non-negative integer $n$, which means 'the reasoner has at least $n$ different beliefs':

$M, s \models min(n)$ iff $|s| \geq n$.

($max(n) \equiv \neg min(n+1)$ means 'at most $n$ different beliefs'.)

# $\Sigma$-bisimulations

Let $M_1 = (S_1, R_1)$ and $M_2 = (S_2, R_2)$ be two structures and $\Sigma \subseteq L$. $M_1$ and $M_2$ are $\Sigma$-**bisimilar** if there is a relation $Z \subseteq S_1 \times S_2$ such that

- if $Z(s,t)$ then $s \cap \Sigma = t \cap \Sigma$
- if $Z(s,t)$ and there is a $s' \in S_1$ such that $R_1(s,s')$, then there is a $t' \in S_2$ such that $R_2(t,t')$ and $Z(s',t')$
- if $Z(s,t)$ and there is a $t' \in S_2$ such that $R_2(t,t')$, then there is a $s' \in S_1$ such that $R_1(s,s')$ and $Z(s',t')$.

# Preservation

Let $\phi \in ML(L)$ and $\Sigma = \{\alpha : B\alpha \in Subf(\phi)\}$. Then $\phi$ is preserved under $\Sigma$-bisimilations: if $Z(s,t)$, then $M_1, s \models \phi$ iff $M_2, t \models \phi$.

However, formulas of $ML(L, \bigtriangledown)$ and $ML(L, min(n))$ are not preserved under $\Sigma$-bisimulations.

# Complete axiomatisation for $min(n)$

| | | |
|---|---|---|
| **MIN0** | $min(0)$ | |
| **MIN1** | $min(n) \rightarrow min(m)$ | $m < n$ |
| **MIN2** | $(B\phi_1 \wedge \cdots \wedge B\phi_n) \rightarrow min(n)$ | $\forall_{i \neq j \in [1,n]} \phi_i \neq \phi_j$ |

# Defining $\bigtriangledown$

We can define believing at most the set of formulas $X$ using $max(|X|)$ and $B$. Below, $\triangle X$ stands for believing at least the formulas in $X$, and $\boxtimes X$ for believing exactly the formulas in $X$.

$$\triangle X \equiv \bigwedge_{\alpha \in X} B\alpha$$

$$\boxtimes X \equiv \triangle X \wedge max(|X|)$$

$$\bigtriangledown X \equiv \bigvee_{Y \subseteq X} \boxtimes Y$$

# Complexity

(assuming unary encoding of $n$ in $min(n)$):

- Adding $min(n)$, for every $n$, to classical propositional logic (where propositional variables are of the form $B\alpha$): satisfiability problem is NP-complete.
- Basic modal logic with $min(n)$ (where propositional variables are of the form $B\alpha$): satisfiability problem PSPACE-complete.

Also, no change in the model-checking complexity.

# Conclusions

- Epistemic logic with purely syntactic beliefs can be used to model-check reasoning resources required to reach some belief
- They are also interesting from a logical point of view.