Dynamic Epistemic Logic

Chapter 2: Multi-Agent S5

Bernhard Nebel   Robert Mattmüller

October 17, 2016

1 Language

When we want to define the basic epistemic language, we need sets of agent symbols and sets of atomic propositions to talk about. Specifically, we have:

- A: a finite set of *agent symbols* (often: a, b, a', a'', $\dots$)
- P: a countable set of *atomic propositions* (often: p, q, p', p'', $\dots$)

Definition 1 (Basic epistemic language). *Let P be a countable set of atomic propositions and A be a finite set of agent symbols. Then the language $\mathcal{L}_K$ is defined by the following BNF:*

$$\phi ::= p \mid \neg p \mid (\phi \wedge \phi) \mid K_a \phi,$$

where $p \in P$ and $a \in A$.

We use some common *abbreviations and conventions:*

- $(\phi \vee \psi) = \neg(\neg\phi \wedge \neg\psi)$
- $(\phi \rightarrow \phi) = (\neg\phi \vee \psi)$
- $(\phi \leftrightarrow \psi) = (\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$
- $\top = p \vee \neg p \quad$ for some $p \in P$
- $\bot = \neg\top$

If there is no risk of confusion, outer parentheses can be omitted.

The only interesting addition compared to propositional logic are the *knowledge modalities* K_a, obviously. The formula $K_a\phi$ is read as "agent a knows that ϕ (is true)". Its dual, $\neg K_a \neg\phi$ is read as "agent a considers ϕ as possible", and is often abbreviated as $\hat{K}_a\phi$. For a group of agents $B \subseteq A$, we write $E_B\phi$ to express that everybody in B knows ϕ. So, $E_B\phi$ can be seen as simply an abbreviation of the formula $\bigwedge_{b \in B} K_b\phi$. Again, we can define its dual, $\hat{E}_B\phi$ as $\neg E_B\neg\phi \equiv \bigvee_{b\in B}\hat{K}_b\phi$, which can be read as "some agent b in B considers ϕ as possible".

Sometimes, when writing *iterated operators*, the following convention comes in handy: if X is a modal operator, then X^n is the n-fold application of X. E. g., $K_a^3\phi$ means $K_aK_aK_a\phi$.

Example 1 (Simplified Hanabi). *In simplified Hanabi, we have* four cards *(*$r1$, $r2$, $g1$, $g2$*),* two players *(*a, b*), and just one card per player. We write* p_c *for the fact that player* p *holds card* c*. Thus, for instance,* a_{r1} *is read as "player* a *has card* $r1$*". Consider the situation where player* a *has card* $r1$ *and player* b *has card* $r2$*. In this situation, all of the following formulas are true:*

- a_{r1} *and* b_{r2},
- $K_a b_{r2}$,
- $K_b a_{r1}$,
- $K_a\neg a_{r2}$ *and* $K_b\neg b_{r1}$
  (Notice that, to arrive at this conclusion, we need to make use of our background theory *that contains assertions such as* $\neg(a_{r1}\wedge b_{r1})$.*),*
- $K_a(K_b a_{r1}\vee K_b a_{g1}\vee K_b a_{g2})$.

2 Semantics

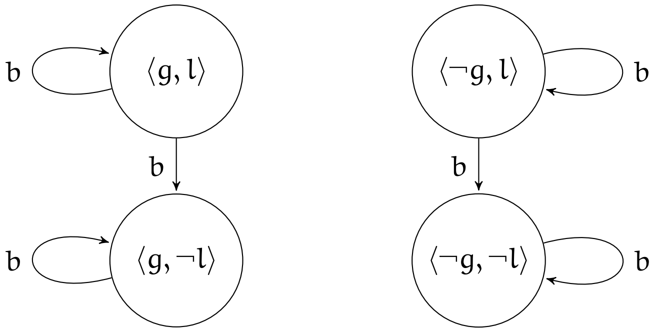
The semantics of the basic epistemic language is based on a special form of "Kripke semantics", where we have *states* (or *worlds*), *accessibility relations* between the worlds, and *propositional valuations* associated with the worlds.

Example 2 (GLO structures). *Consider three cities, namely Groningen, Liverpool and Otago. Assume that:*

- *Person* b *lives in Groningen.*
- *Person* w *lives in Liverpool.*
- *"The weather in Groningen is sunny" is the atomic proposition* g.
- *"The weather in Liverpool is sunny" is the atomic proposition* l.

States are just possible weather conditions: $\langle g,l\rangle$, $\langle\neg g,l\rangle$, $\langle g,\neg l\rangle$, $\langle\neg g,\neg l\rangle$*. We want to model what agent* b *knows. Assume that* b *is in state* $\langle g,l\rangle$*. He also considers the state* $\langle g,\neg l\rangle$ *possible.*

This situation can be graphically captured by the following model M_1*:*



This is an example of a so-called Kripke model.

Definition 2 (Kripke model). *Given a countable set of atomic propositions* P *and a finite set of agent names* A*, a* Kripke model *is a structure* $M = (S, R_A, V_P)$ *where:*

- S *is a set of states (also called the* domain *of* M*, in symbols* $D(M)$*),*
- R_A *is a function yielding for every* $a \in A$ *an* accessibility relation $R_A(a) = R_a \subseteq S \times S$.
- $V_P : P \to 2^S$ *is a* valuation function *that for all* $p \in P$ *yields the set of worlds* $V_P(p) \subseteq S$ *where* p *is true.*

If A and P are not important or clear from the context, we will often drop them and write $M = (S, R, V)$. If all accessibility relations R_a are equivalence relations (reflexive, symmetric and transitive), then we also use the symbols $\sim$ for R and $\sim_a$ for R_a. In that case, $M = (S, \sim, V)$ is also called an *epistemic model.*

Formulas are then interpreted over states in models (aka. states, pointed models, epistemic states).

Example 3. *Assume we have the formula* $K_b l$*. This formula is* not *true in state* $\langle \neg g, l \rangle$*, symbolically* $\langle \neg g, l \rangle \not\models K_b l$*, because in* $\langle \neg g, l \rangle$*, agent* b *also considers world* $\langle \neg g, \neg l \rangle$ *possible, and in that world,* l *does not hold.*

We can define truth of an epistemic formula in an epistemic state inductively as follows.

Definition 3. *Given a Kripke model* $M = (S, R, V)$*, the pair* (M, s) *is called a* pointed model *for* $s \in S$*. If* M *is an epistemic model, then* (M, s) *is called an* epistemic state.

Definition 4. *A formula* ϕ *is true in an epistemic state* (M, s)*, symbolically* $M, s \models \phi$*, under the following conditions:*

- $M, s \models p$ *iff* $s \in V(p)$.
- $M, s \models \phi \wedge \psi$ *iff* $M, s \models \phi$ *and* $M, s \models \psi$.

- $M, s \models \neg\phi$ *iff* $M, s \not\models \phi$.
- $M, s \models K_a\phi$ *iff* $M, t \models \phi$ *for all* $t \in S$ *such that* $(s, t) \in R_a$.

This implies, among others, that $M, s \models \hat{K}_a\phi$ iff $M, t \models \phi$ for *some* $t \in S$ such that $(s, t) \in R_a$.

Definition 5. *If (M,s)* $\models \phi \forall s \in D(M)$ *then we say that* ϕ *is true in M*

Definition 6. *If M* $\models \phi$ *for all models in M in a certain class* $\mathcal{X}$ *of models then we say* ϕ *is valid in* $\mathcal{X}$*, symbollicaly* $\mathcal{X} \models \phi(L)$

Definition 7. *If for the formula* ϕ *there exists a pointed model (M,s) such that* ϕ *is true in this pointed model then we say* ϕ *is satisfied in M.*
If M belongs to a class of models $\mathcal{X}$ *then* ϕ *is satisfiable in* $\mathcal{X}$.

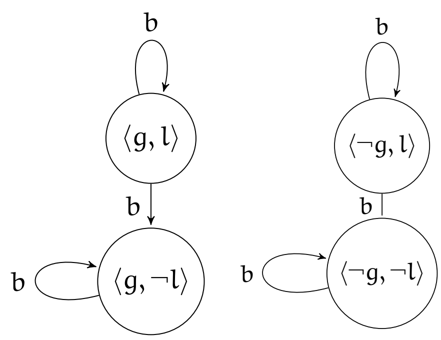


Figure 1: Model M1

Example 4. *Let us check if* $(M_1, \langle g, l\rangle) \models \hat{K}_b g \wedge \hat{K}_b l \wedge \hat{K}_b \neg l$
We can see that $\hat{K}_b g$ *is true because of the reflexive relation. The same can be said for* $\hat{K}_b l$ *for the same reason.* $\hat{K}_b \neg l$ *is also true. Thus the whole formula is true.*

Example 5. $(M_1, \langle g, l\rangle) \models K_b(K_b g and \neg K_b l)$.
We have to go to all states and see if the formula inside holds. $K_b g$ *is true because in all states g is true. The second part is true because there is a state, namely* $\langle g, \neg l\rangle$ *where l is not true.*

Example 6. $M_1 \models (K_g \vee K_b \neg g) \wedge (\neg K_b \wedge \neg K_b \neg l)$
It is easy to see that both clauses are true and thus the whole formula is true.

Convention:
Visualizations of epistemic models use undirected edges and leave out the reflexive (and sometimes also transitive edges).

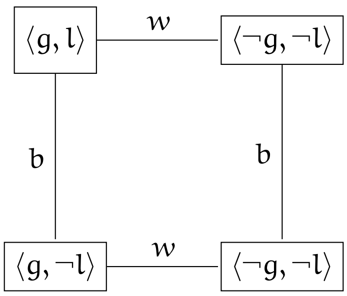


Figure 2: Model M_2

Example 7. $(M_2, \langle g, l\rangle) \models (K_b g \lor K_b \neg g) \land (K_w l \lor K_w \neg l)$ *From the graphical representation of M_2 we can see that the formula above holds in* $(M_2, \langle g, l\rangle)$.

Can we come up with a model that has a bit of ignorance about what the other agent knows. To do that we need to introduce more worlds. Note that there can be distinct states with identical valuations!

Example 8. *Another agent h (from Otago) calls w. w tells h that l is true. h tells w that he will call b afterwards (but he does not say whether he will tell b about l). So w does not know whether b knows that l is true.*

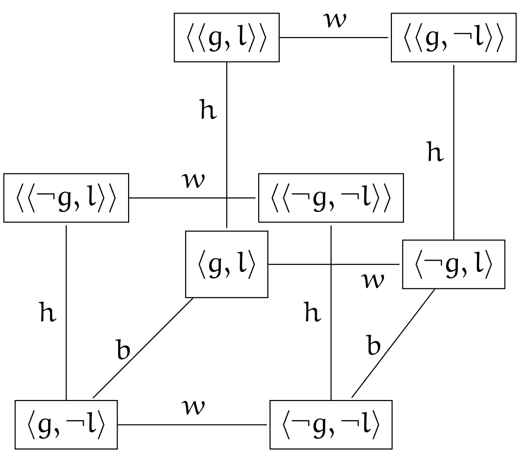


Figure 3: Model M_3

Example 9. $(M_3, \langle g, l\rangle) \models l \land \neg K_b l \land K_b(\neg K_w K_b l \land \neg K_w \neg K_b l)$

Proposition:

Let ϕ and ψ be formulas of our language $\mathcal{L}_K$ and let K_a be an epistemic operator for some $a \in A$. Let $\mathcal{K}$ be the set of all Kripke models and $\mathcal{S}5$ be the set of all epistemic models. Then the following hold:

(LO1) $\mathcal{K} \models K_a\phi \wedge K_a(\phi \rightarrow \psi) \rightarrow K_a\psi$
(LO2) $\mathcal{K} \models \phi$ implies $\mathcal{K} \models K_a\phi$
(LO3) $\mathcal{K} \models \phi \rightarrow \psi$ implies $\mathcal{K} \models K_a\phi \rightarrow K_a\psi$
(LO4) $\mathcal{K} \models \phi \leftrightarrow \psi$ implies $\mathcal{K} \models K_a\phi \leftrightarrow K_a\psi$
(LO5) $\mathcal{K} \models (K_a\phi \wedge K_a\psi) \rightarrow K_a(\phi \wedge \psi)$
(LO6) $\mathcal{K} \models K_a\phi \rightarrow K_a(\phi \vee \psi)$
(LO7) $\mathcal{S}5 \models \neg(K_a\phi \wedge K_a\neg\phi)$

Definition 8. *Relation properties*
A relation R is called reflexive if $\forall s$ *we have* $(s, s) \in R$.
A relation R is called symmetric $\forall s, t$ *we have if* $(s, t) \in R$ *implies* $(t, s) \in R$.
A relation R is called transitive if $\forall s, t, u$ *we have* $(s, t) \in R$, $(t, u) \in R$ *then* $(s, u) \in R$.
A relation R is called serial if $\forall s, \exists t$ *s.t* $(s, t) \in R$.
A relation is called Eucledian if $\forall s, t, u$*: if* $(s, t) and (s, u) \in R$ *then* $(t, u) \in R$.

Definition 9. *Modal classes*
Kripke models are classified according to the properties of the accessibility relation R_a *as follows:*

Relation property	Name
No restriction	$\mathcal{K}$
Serial	$\mathcal{KD}$
Reflexive	$\mathcal{T}$
Transitive	$\mathcal{K}4$
Reflexive and transitive	$\mathcal{K}5$
Transitive and euclidian	$\mathcal{K}45$
Serial, transitive and euclidian	$\mathcal{KD}45$
Serial, transitive, euclidian and reflexive	$\mathcal{S}5$

Definition 10. *Bisimulation*
Let two M=(S,R,V) and M'=(S',R',V') be given. A non-empty relations $\mathcal{B} \subseteq S \times S$ *is a* <u>*bisimulation*</u> *iff* $\forall s \in S$, $s' \in S'$ *with* $(s, s') \in \mathcal{B}$ *it holds:*
(atoms) $s \in V(p)$ *iff* $s' \in V'(p) \forall p \in P$.
(forth) $\forall a \in A$ *and all states* $t \in S$ *if* $(s, t) \in R_a$ *then* $\exists t' \in S'$ *s.t* $(s', t') \in R'_a$ *and also* $(t, t') \in \mathcal{B}$.
(back) $\forall a \in A$ *and* $\forall$ *t'* $\in S'$, *if we have* $(s', t') \in R'_a$ *then* $\exists t \in S$ *s.t* $(s, t) \in R_a$ *and* $(t, t') \in \mathcal{B}$.

We write (M,s) $\underline{\leftrightarrow}$ (M',s') iff there is a bisimulation and we say (M,s) and (M',s') are bisimilar.
Bisimilarity means that you cannot distinguish between the states that are bisimilar.
We write $(M, s) \equiv_{\mathcal{L}_K} (M', s')$ iff $(M, s) \models \phi$ iff $(M', s') \models \phi \ \forall \phi \in \mathcal{L}_K$

Theorem 1. *Bisimulation*
For all pointed models (M,s) and (M',s'): (M,s) $\underline{\leftrightarrow}$ *(M',s') then* $(M, s) \equiv_{\mathcal{L}_K} (M', s')$

Proof(by structural induction):
Base case: Suppose (M,s) $\underline{\leftrightarrow}$ (M',s').
By the atoms condition it must be the case that $(M, s) \models p$ iff $(M', s') \models p \ \forall p \in P$.
Induction hypothesis: (M,s) $\underline{\leftrightarrow}$ (M',s') implies $(M, s) \models \phi$ iff $(M', s') \models \phi$.
Induction step:
Negation: Suppose that $(M, s) \models \neg\phi$. By the definition of negation it means $(M, s) \not\models \phi$. By the induction hypothesis we know that this is equivalent to $(M', s') \not\models \phi$ i.e. $(M', s') \models \neg\phi$.
Conjunction: Suppose $(M, s) \models \phi \wedge \psi$, i.e. $(M, s) \models \phi$ and $(M, s) \models \psi$. By applying the I.H. we have that $(M', s') \models \phi$ and $(M', s') \models \psi$, i.e. $(M', s') \models \phi \wedge \psi$.
Epistemic operator: We suppose that $(M, s) \models K_a\phi$. We consider an arbitrary t' with the property $(s', t') \in R'_a$. By the back property of bisimulation there exists a t s.t. $(s.t) \in R_a$ and $(t, t') \in \mathcal{B}$. By I.H. we know that $(M, t) \models \phi$ iff $(M', t') \models \phi$. Since $(M, s) \models K_a\phi$ and $(M, t) \models \phi$ that means that $(M, t') \models \phi$. Since t' is an arbitrary choice, this hold for all t' s.t. $(s', t') \in R'_a$, therefore $(M', s') \models K_a\phi$.
The other way around is analogous by using forth instead of back.

Note: Bisimulation implies equivalence of the set of two formulas, but the converse does not hold.
Note: The proof applies to all classes of models.