Justification Logic

Thomas Studer

Institute of Computer Science
University of Bern
Bern, Switzerland

Modal Logic

Modal logic adds a new connective $\Box$ to the language of logic.

Two traditions:

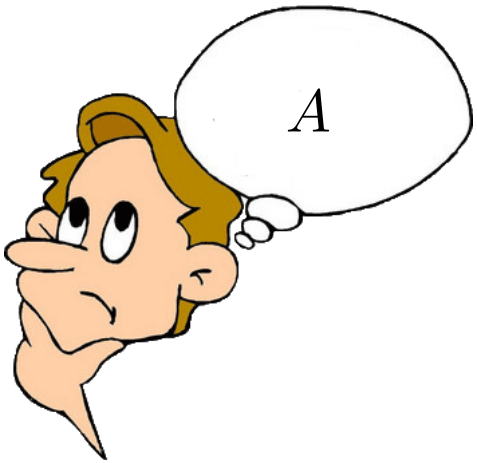
Epistemic logic:

$\Box A$ means A is known / believed

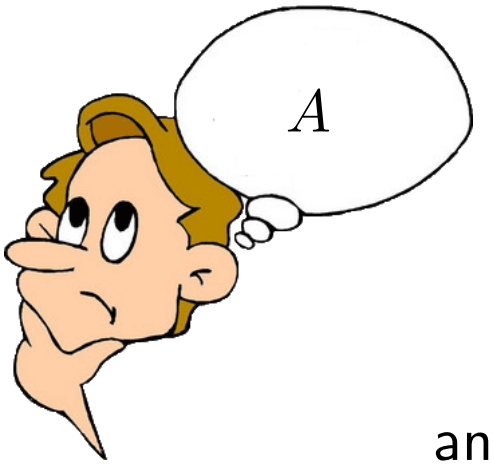
Proof theory:

$\Box A$ means A is provable in system S

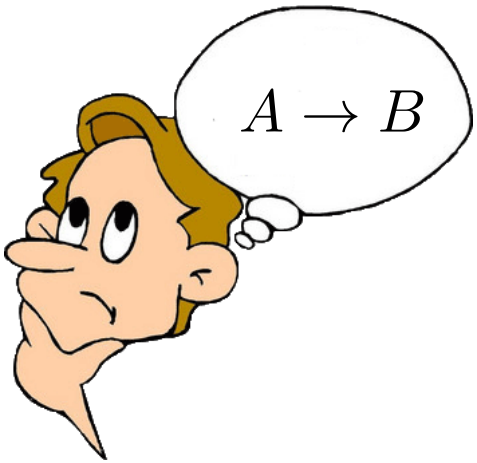
Modal Logic: How It Works



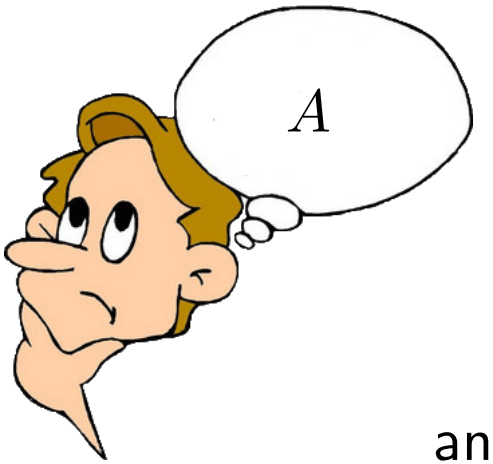
Modal Logic: How It Works



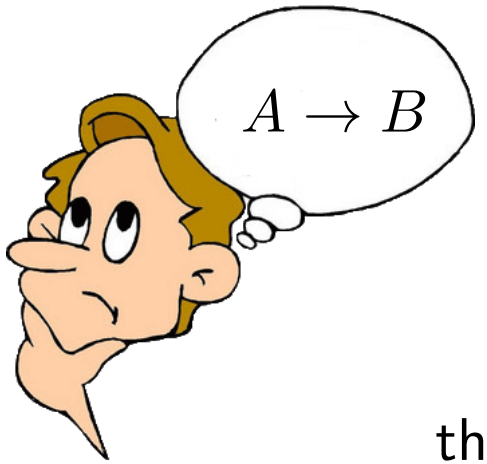
and



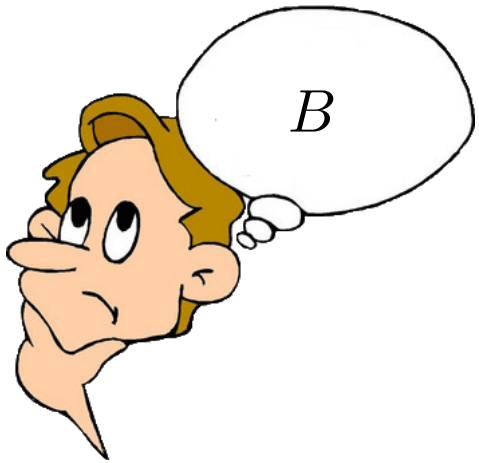
Modal Logic: How It Works



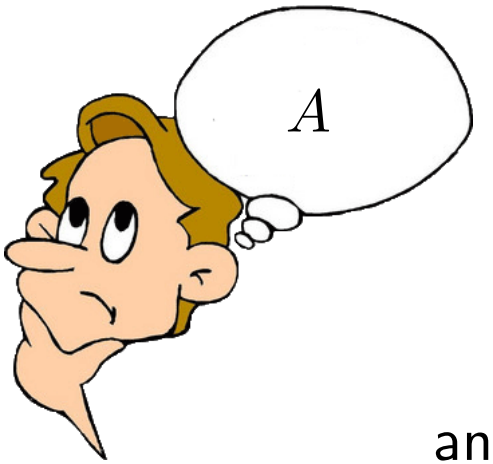
and



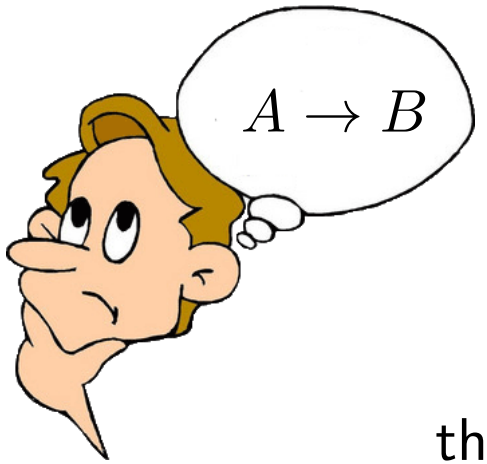
thus



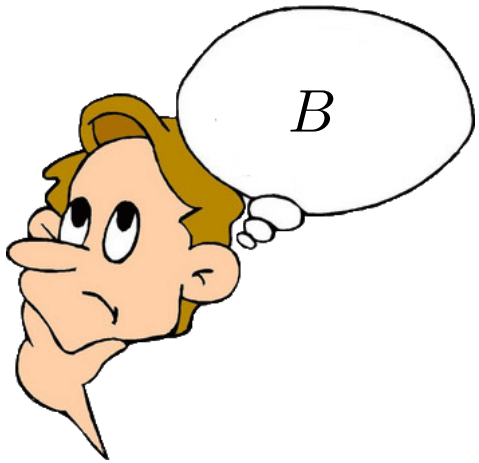
Modal Logic: How It Works



and

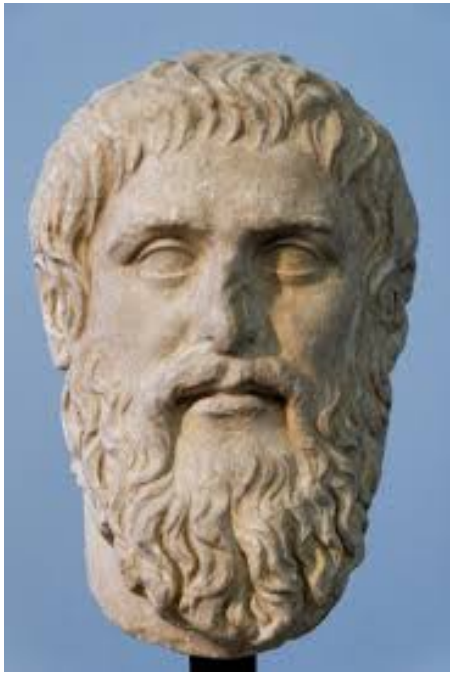


thus



$$\Box A \quad \wedge \quad \Box(A \to B) \quad \to \quad \Box B$$

Problems: Epistemic Tradition



Plato:

Knowledge is justified true belief

True belief is modeled by $\Box A \rightarrow A$ but

where are the justifications in modal logic?

Problems: Proof-Theoretic Tradition

$\Box\bot \rightarrow \bot$ is an axiom, which means

$\neg\Box\bot$ is provable. Hence, by necessitation

$\Box\neg\Box\bot$ is provable.

Problems: Proof-Theoretic Tradition

$\Box\bot \rightarrow \bot$ is an axiom, which means

$\neg\Box\bot$ is provable. Hence, by necessitation

$\Box\neg\Box\bot$ is provable.

$\Box\bot$ means S proves $\bot$.

$\neg\Box\bot$ means S does not prove $\bot$, that is

$\neg\Box\bot$ means S is consistent.

$\Box\neg\Box\bot$ means S proves that S is consistent.

Problems: Proof-Theoretic Tradition

$\Box\bot \rightarrow \bot$ is an axiom, which means

$\neg\Box\bot$ is provable. Hence, by necessitation

$\Box\neg\Box\bot$ is provable.

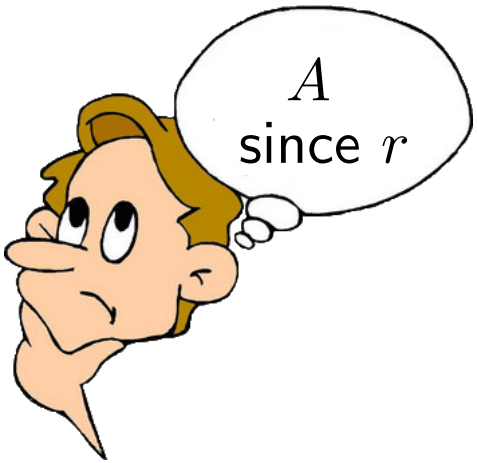
$\Box\bot$ means S proves $\bot$.

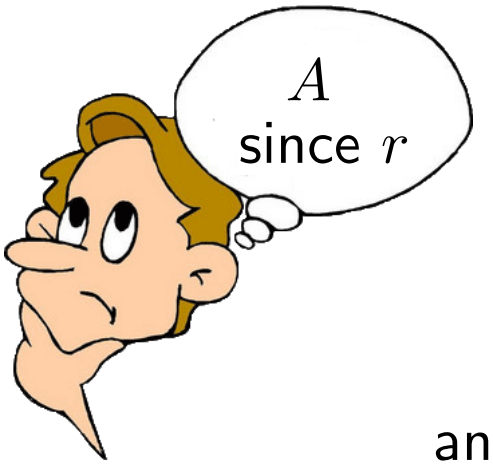
$\neg\Box\bot$ means S does not prove $\bot$, that is

$\neg\Box\bot$ means S is consistent.

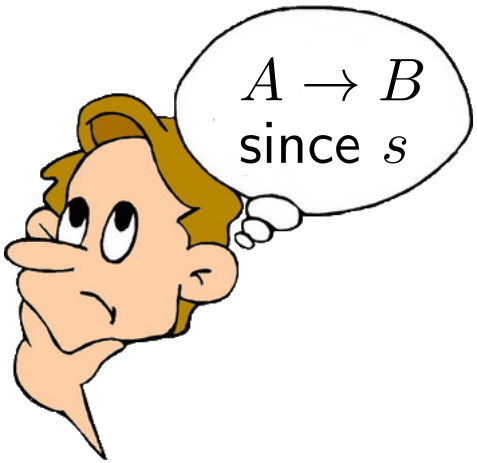
$\Box\neg\Box\bot$ means S proves that S is consistent.

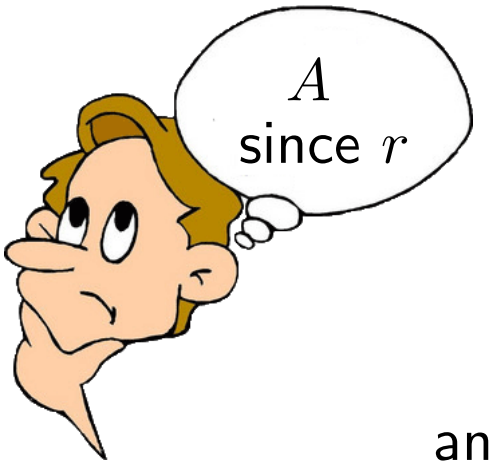
Gödel: if S has a certain strength, it cannot prove its own consistency.


A
since r

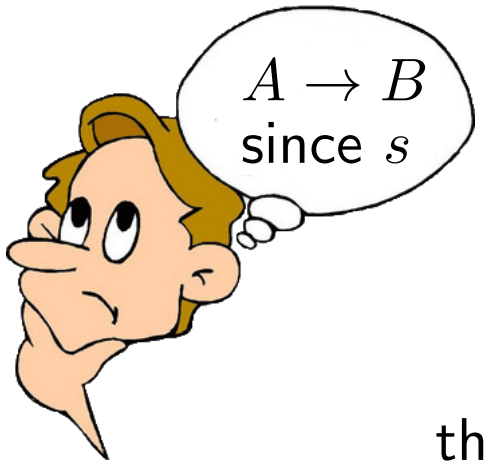


and

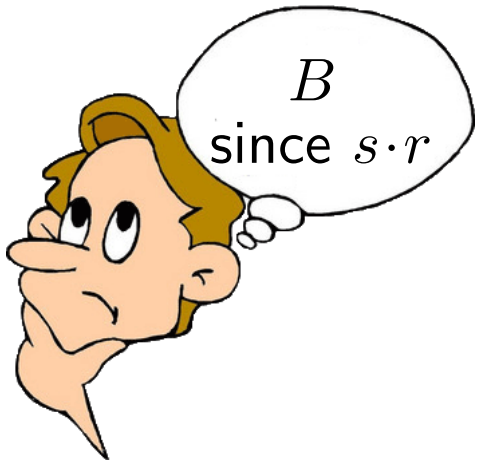


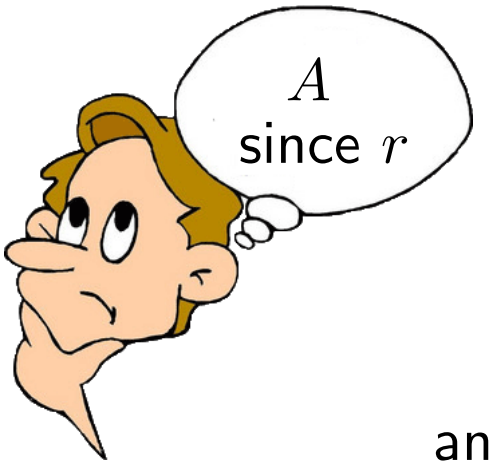


and

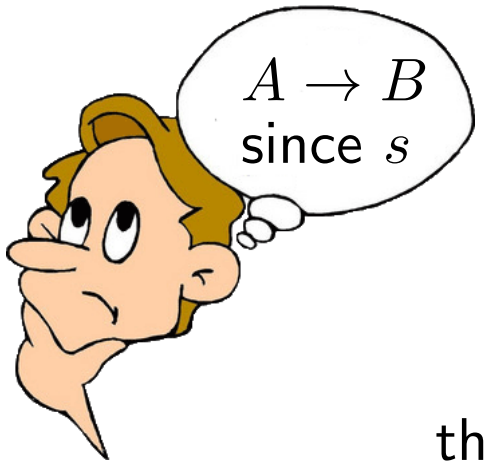


thus

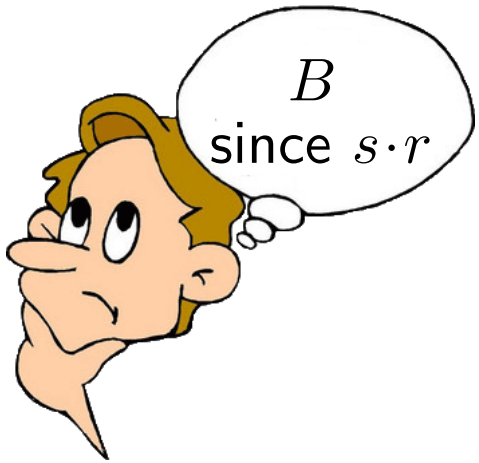




and



thus



$$r{:}A \quad \wedge \quad s{:}(A \to B) \quad \to \quad s \cdot r{:}B$$

Syntax of the Logic of Proofs

Logic

The logic of proofs $\mathsf{LP}_{\mathsf{CS}}$ is the justification counterpart of the modal logic S4.

Justification terms Tm

$$t ::= x \mid c \mid (t \cdot t) \mid (t + t) \mid !t$$

Formulas $\mathcal{L}_j$

$$A ::= p \mid \neg A \mid (A \to A) \mid t{:}A$$

Axioms for LP

- all propositional tautologies
- $t{:}(A \to B) \to (s{:}A \to (t \cdot s){:}B)$ (J)
- $t{:}A \to (t + s){:}A, \quad s{:}A \to (t + s){:}A$ (+)
- $t{:}A \to A$ (jt)
- $t{:}A \to !t{:}t{:}A$ (j4)

Deductive System

Constant specification

A constant specification CS is any subset

$$\mathsf{CS} \subseteq \{(c, A) \mid c \text{ is a constant and } A \text{ is an axiom}\}.$$

The deductive system $\mathsf{LP}_{\mathsf{CS}}$ consists of the above axioms and the rules of modus ponens and axiom necessitation.

$$\frac{A \quad A \to B}{B} \qquad\qquad \frac{(c, A) \in \mathsf{CS}}{c{:}A}$$

Necessitation

Definition

A constant specification CS for LP is called *axiomatically appropriate* if for each axiom F of LP, there is a constant c such that $(c, F) \in \mathsf{CS}$.

Lemma (Constructive Necessitation)

Let CS *be an axiomatically appropriate constant specification. For any formula* A*, if*

$$\mathsf{LP}_{\mathsf{CS}} \vdash A,$$

then

$$\mathsf{LP}_{\mathsf{CS}} \vdash t{:}A$$

for some term t*.*

Forgetful Projection

Definition (Forgetful projection)

The mapping $\circ$ from justified formulas to modal formulas is defined as follows

1. $P^\circ := P$ for P atomic;
2. $(\neg A)^\circ := \neg A^\circ$;
3. $(A \to B)^\circ := A^\circ \to B^\circ$;
4. $(t{:}A)^\circ := \Box A^\circ$.

Lemma (Forgetful projection)

For any constant specification CS *and any formula* F *we have*

$$\mathsf{LP}_{\mathsf{CS}} \vdash F \quad \textit{implies} \quad \mathsf{S4} \vdash F^\circ \ .$$

Realization

Definition (Realization)

A *realization* is a mapping r from modal formulas to justified formulas such that $(r(A))^\circ = A$.

Definition

We say a justification logic $\mathsf{LP}_{\mathsf{CS}}$ *realizes* S4 if there is a realization r such that for any formula A we have

$$\mathsf{S4} \vdash A \quad \text{implies} \quad \mathsf{LP}_{\mathsf{CS}} \vdash r(A) \ .$$

Realization Theorem

Definition (Schematic CS)

We say that a constant specification is *schematic* if it satisfies the following: for each constant c, the set of axioms $\{A \mid (c, A) \in \mathsf{CS}\}$ consists of all instances of one or several (possibly zero) axiom schemes of LP.

Theorem (Realization)

Let CS *be an axiomatically appropriate and schematic constant specification. There exists a realization* r *such that for all formulas* A

$$\mathsf{S4} \vdash A \quad \Longrightarrow \quad \mathsf{LP}_{\mathsf{CS}} \vdash r(A) \ .$$

Self-referentiality

Definition (Self-referential CS)

A constant specification CS is called *self-referential* if $(c, A) \in \mathsf{CS}$ for some axiom A that contains at least one occurrence of the constant c.

S4 and $\mathsf{LP}_{\mathsf{CS}}$ describe self-referential knowledge. That means if $\mathsf{LP}_{\mathsf{CS}}$ realizes S4 for some constant specification CS, then that constant specification must be self-referential.

Lemma

Consider the S4*-theorem* $G := \neg\Box((P \to \Box P) \to \bot)$ *and let* F *be any realization of* G*.*
If $\mathsf{LP}_{\mathsf{CS}} \vdash F$*, then* CS *must be self-referential.*

Arithmetical Semantics

Originally, $\mathsf{LP}_{\mathsf{CS}}$ was developed to provide classical provability semantics for intuitionistic logic.

Arithmetical Semantics for $\mathsf{LP}_{\mathsf{CS}}$: Justification terms are interpreted as proofs in Peano arithmetic and operations on terms correspond to computable operations on proofs in PA.

$$\mathsf{Int} \xrightarrow{\text{Gödel}} \mathsf{S4} \xrightarrow{\text{Realization}} \mathsf{JL} \xrightarrow{\text{Arithm. sem.}} \mathsf{CL} + \text{proofs}$$

Basic evaluation

Definition (Basic Evaluation)

A *basic evaluation* $*$ for $\mathsf{LP}_{\mathsf{CS}}$ is a function:

$$* : \mathsf{Prop} \to \{0,1\} \text{ and } * : \mathsf{Tm} \to \mathcal{P}(\mathcal{L}_j), \text{ such that}$$

1. $F \in (s \cdot t)^*$ if $(G \to F) \in s^*$ and $G \in t^*$ for some G
2. $F \in (s + t)^*$ if $F \in s^*$ or $F \in t^*$
3. $F \in t^*$ if $(t, F) \in \mathsf{CS}$
4. $s{:}F \in (!s)^*$ if $F \in s^*$

Quasimodel

Definition (Quasimodel)

A *quasimodel* is a tuple $\mathcal{M} = (W, R, *)$ where $W \neq \varnothing$, $R \subseteq W \times W$, and the *evaluation* $*$ maps each world $w \in W$ to a basic evaluation $*_w$.

Definition (Truth in quasimodels)

$\mathcal{M}, w \Vdash p$ if and only if $p^*_w = 1$ for $p \in$ Prop;

$\mathcal{M}, w \Vdash F \to G$ if and only if $\mathcal{M}, w \nVdash F$ or $\mathcal{M}, w \Vdash G$;

$\mathcal{M}, w \Vdash \neg F$ if and only if $\mathcal{M}, w \nVdash F$;

$\mathcal{M}, w \Vdash t{:}F$ if and only if $F \in t^*_w$.

Model

Given $\mathcal{M} = (W, R, *)$ and $w \in W$, we define

$$\Box_w := \{F \in \mathcal{L}_j \mid \mathcal{M}, v \Vdash F \text{ whenever } R(w, v)\} \ .$$

Definition (Modular Model)

A *modular model* $\mathcal{M} = (W, R, *)$ is a quasimodel with

1. $t^*_w \subseteq \Box_w$ for all $t \in \mathsf{Tm}$ and $w \in W$; (JYB)
2. R is reflexive;
3. R is transitive.

Theorem (Soundness and Completeness)

For all formulas $F \in \mathcal{L}_j$,

$$\mathsf{LP}_{\mathsf{CS}} \vdash F \quad \Longleftrightarrow \quad \mathcal{M} \Vdash F \text{ for all modular models } \mathcal{M}.$$

Decidability

In modal logic, decidability is a consequence of the finite model property. For $\mathsf{LP}_{\mathsf{CS}}$ the situation is more complicated since CS usually is infinite.

Theorem

$\mathsf{LP}_{\mathsf{CS}}$ *is decidable for decidable schematic constant specifications* CS*.*

A decidable CS is not enough:

Theorem

There exists a decidable constant specification CS *such that* $\mathsf{LP}_{\mathsf{CS}}$ *is undecidable.*

Complexity

Theorem

Let CS *be a schematic constant specification.*
The problem whether $\mathsf{LP}_{\mathsf{CS}} \vdash t{:}B$ *belongs to NP.*

Complexity

Theorem

Let CS *be a schematic constant specification.*
The problem whether $\mathsf{LP}_{\mathsf{CS}} \vdash t{:}B$ *belongs to NP.*

Definition

A constant specification is called *schematically injective* if it is schematic and each constant justifies no more that one axiom scheme.

Theorem

Let CS *be a schematically injective and axiomatically appropriate constant specification.*
The derivability problem for $\mathsf{LP}_{\mathsf{CS}}$ *is* Π_2^p*-complete.*

Logical Omniscience

Modal logic of knowledge contains the epistemic closure principle in the form of axiom

$$\Box(A \to B) \to (\Box A \to \Box B) \ ,$$

which yields an unrealistic feature called *logical omniscience* whereby an agent knows all logical consequences of her assumptions.

Definition

A *proof system* for a logic L is a binary relation $E \subset \Sigma^\star \times \mathsf{L}$ between words in some alphabet, called proofs, and theorems of L such that

1. E is computable in polynomial time and
2. for all formulas F, $\mathsf{L} \vdash F$ if and only if there exists y with $E(y, F)$.

Logical Omniscience II

Knowledge assertion A is a provable formula of the form

$$\Box B \text{ for S4} \qquad \text{or} \qquad t{:}B \text{ for } \mathsf{LP}_{\mathsf{CS}}$$

with the object of knowledge function $\mathsf{OK}(A) := B$.

Definition (Logical Omniscience Test (LOT))

An proof system E for an epistemic logic L is *not logically omniscient*, or *passes LOT*, if there exists a polynomial P such that for any knowledge assertion A, there is a proof of $\mathsf{OK}(A)$ in E with the size bounded by $P(\mathsf{size}(A))$.

Logical Omniscience III

Theorem (S4 is logically omniscient)

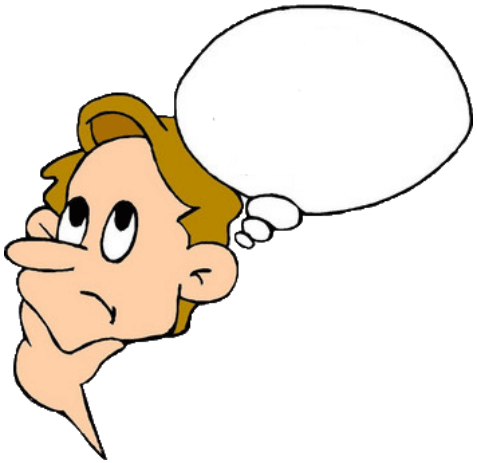
There is no proof system for S4 *that passes LOT unless NP=PSPACE.*

Theorem ($\mathsf{LP}_{\mathsf{CS}}$ is not logically omniscient)

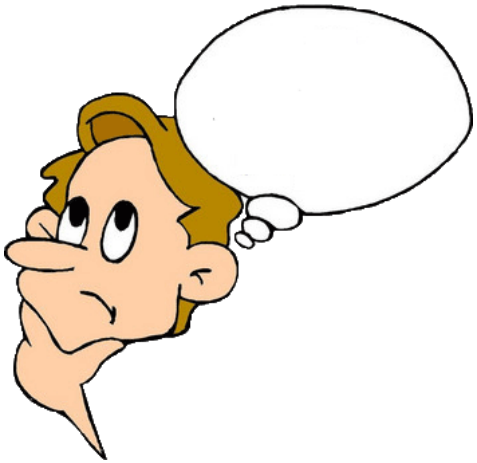
Let CS *be a schematic constant specification. There is a proof system for* $\mathsf{LP}_{\mathsf{CS}}$ *that passes LOT.*

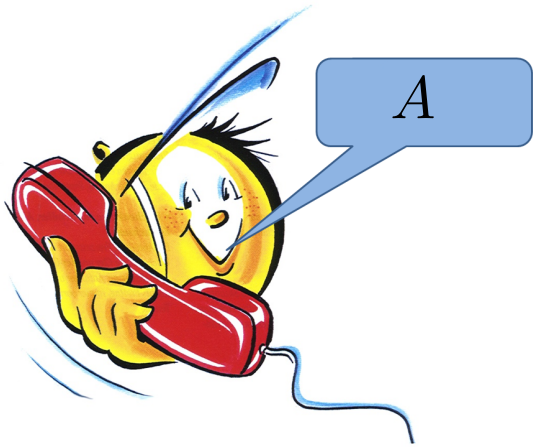
Dynamic Epistemic Logic

Dynamic Epistemic Logic

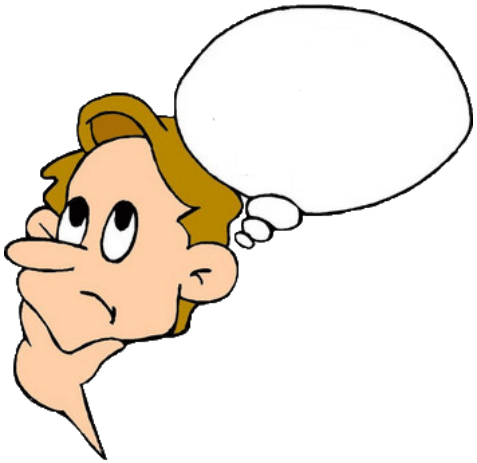


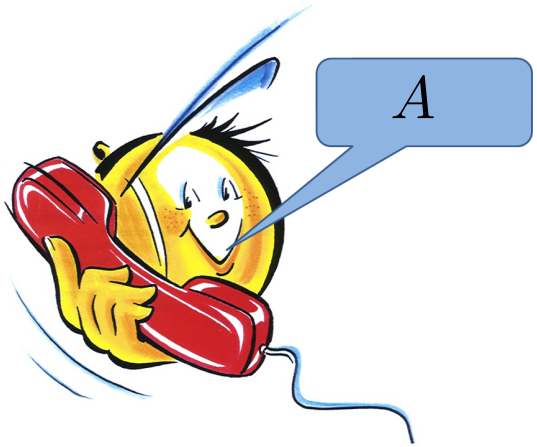
Dynamic Epistemic Logic

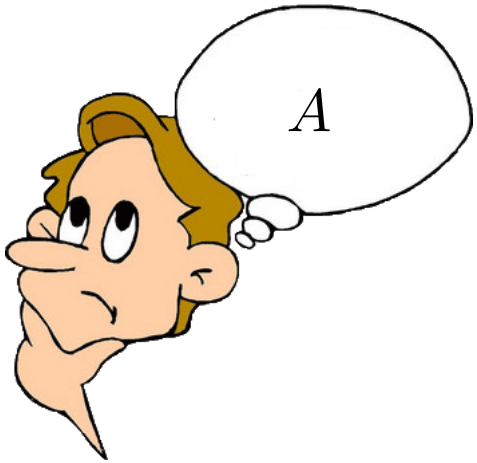




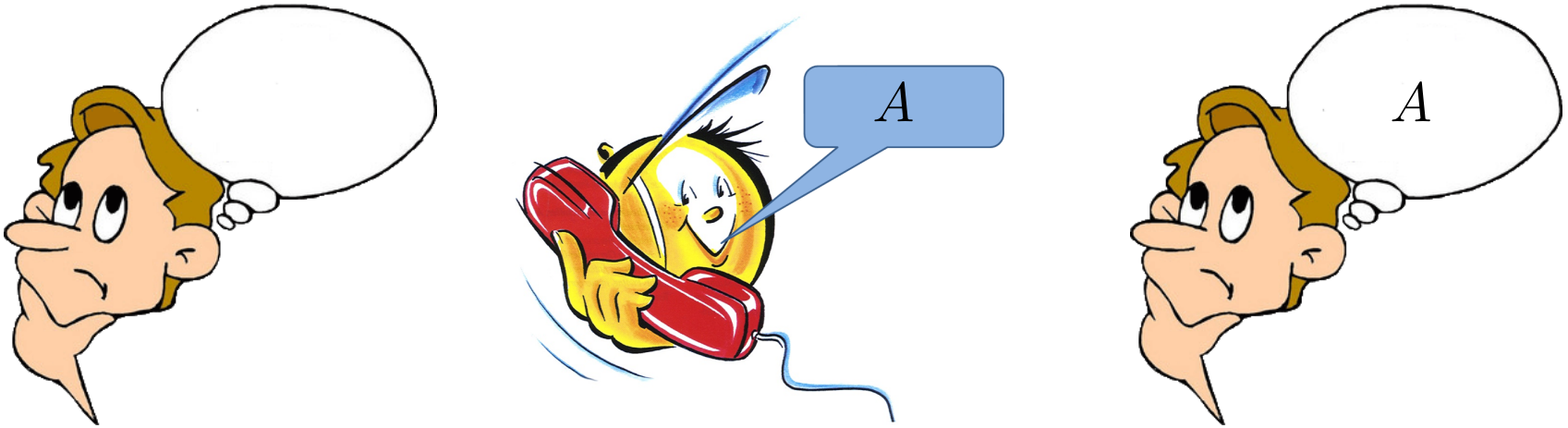
Dynamic Epistemic Logic





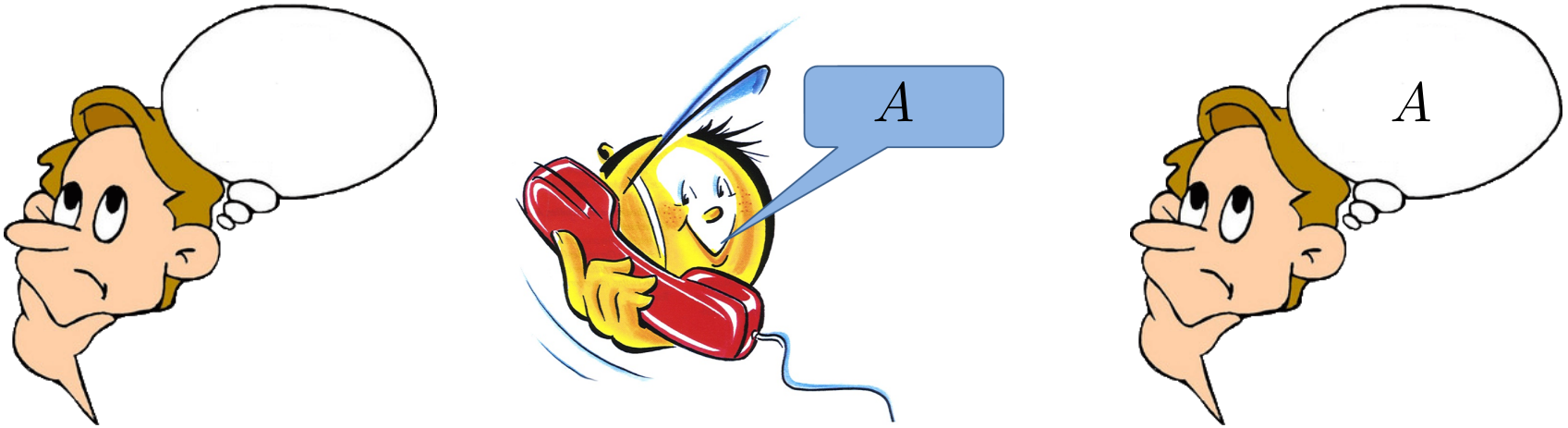


Dynamic Epistemic Logic



After the announcement of A, the agent believes A, i.e. $[A]\Box A$

Dynamic Epistemic Logic



After the announcement of A, the agent believes A, i.e. $[A]\Box A$

Problem

The $\Box$-operator does not tell us whether A is believed because of the announcement or whether A is believed independent of it.

Update as Evidence: the Logic $\mathsf{JUP}_{\mathsf{CS}}$ for Belief Expansion

Fundamental principle

After the announcement of A,

the announcement itself justifies the agent's belief in A.

Update as Evidence: the Logic $\mathsf{JUP}_{\mathsf{CS}}$ for Belief Expansion

Fundamental principle

After the announcement of A,

the announcement itself justifies the agent's belief in A.

For each formula A we add a new justification term $\mathsf{up}(A)$.

Some axioms of JUP:

- Success: $[A]\,\mathsf{up}(A){:}A$

Update as Evidence: the Logic $\mathsf{JUP}_{\mathsf{CS}}$ for Belief Expansion

Fundamental principle

After the announcement of A,

the announcement itself justifies the agent's belief in A.

For each formula A we add a new justification term $\mathsf{up}(A)$.

Some axioms of JUP:

- Success: $[A]\,\mathsf{up}(A){:}A$
- Persistence: $t{:}B \rightarrow [A]t{:}B$.
- Reduction axioms
- Minimal change
- Iterated updates

Basic Properties of $\mathsf{JUP}_{\mathsf{CS}}$

Lemma (Minimal change)

Let t be a term that does not contain $\mathsf{up}(A)$ *as a subterm. Then*

$$\mathsf{JUP}_{\mathsf{CS}} \vdash [A]t{:}B \leftrightarrow t{:}B \quad .$$

Basic Properties of $\mathsf{JUP}_{\mathsf{CS}}$

Lemma (Minimal change)

Let t be a term that does not contain $\mathsf{up}(A)$ as a subterm. Then

$$\mathsf{JUP}_{\mathsf{CS}} \vdash [A]t{:}B \leftrightarrow t{:}B \ .$$

Lemma (Ramsey I)

$\mathsf{JUP}_{\mathsf{CS}} \vdash t{:}(A \to B) \to [A](t{\cdot}\mathsf{up}(A)){:}B.$

Lemma (Ramsey II)

Let CS be an axiomatically appropriate constant specification. For each term t there exists a term s such that

$$\mathsf{JUP}_{\mathsf{CS}} \vdash [A]t{:}B \to s{:}(A \to B) \ .$$

Thank you!

A Justified Version of $\Box A \vee \Box B \to \Box(A \vee B)$

Assume we are given $\mathsf{LP}_{\mathsf{CS}}$ with

$$(a, A \to (A \vee B)) \in \mathsf{CS} \quad \text{and} \quad (b, B \to (A \vee B)) \in \mathsf{CS} \ .$$

By axiom necessitation we get

$$\mathsf{LP}_{\mathsf{CS}} \vdash a{:}(A \to (A \vee B)) \quad \text{and} \quad \mathsf{LP}_{\mathsf{CS}} \vdash b{:}(B \to (A \vee B)) \ .$$

Using (J) and (MP) we obtain

$$\mathsf{LP}_{\mathsf{CS}} \vdash x{:}A \to (a \cdot x){:}(A \vee B) \quad \text{and} \quad \mathsf{LP}_{\mathsf{CS}} \vdash y{:}B \to (b \cdot y){:}(A \vee B) \ .$$

Finally, from $(+)$ we have

$$\mathsf{LP}_{\mathsf{CS}} \vdash (a \cdot x){:}(A \vee B) \to (a \cdot x + b \cdot y){:}(A \vee B) \text{ and}$$
$$\mathsf{LP}_{\mathsf{CS}} \vdash (b \cdot y){:}(A \vee B) \to (a \cdot x + b \cdot y){:}(A \vee B) \ .$$

Using propositional reasoning, we obtain

$$\mathsf{LP}_{\mathsf{CS}} \vdash (x{:}A \vee y{:}B) \to (a \cdot x + b \cdot y){:}(A \vee B) \ .$$