LFAI 2002:

Logical Foundations
of Artificial Intelligence
(*Lecture 7*)

Sergei N. Artemov

April 3, 2002

Computer Science Program
CUNY Graduate Center

This lecture in brief

1. Modal logic: time and knowledge
2. Basic systems of modal logic
3. Possible worlds semantics
4. Logics of linear time
5. Logics of branching time
6. Temporal logic and verification

Modal logic: time and knowledge

Propositional logic is decidable but too restrictive. First order and higher order logics have unlimited expressive power but are not decidable. Modal logic appeared as an attempt to extend propositional logic by additional connectives preserving certain nice features like decidability.

Minimal format: propositional connectives plus unary connective "modality" $\Box$. Intended readings of new atoms $\Box F$ are
1. Epistemic - existential: "F is known", "F is provable", etc,
2. Temporal - universal: "F holds in all possible situations", "in the future F will always hold", etc.

Usually preserves decidability.

History and Applications

McKinsey-Tarski (1948): topological semantics $\Box F = \mathit{interior}(F)$, provides a mathematical model for intuitionism, logic of approximate measurements, leads to logics for dynamic systems, etc.

Kripke (1959): possible worlds à la Leibniz, by far the most widely used semantics.

Hoare (1969): partial correctness statements $A\{G\}B =$"if A holds before the execution of G then B holds afterward", a classic of program verification. Recently Tony Hoare was knighted by the British Queen.

Pratt (1976): logic of programs, $[C]\varphi = \varphi$ holds while C is executed, each $[C]$ is an **S4**-modality. Kripke style semantics where possible worlds are machine states. Stanford University Network = (SUN).

Pnueli (1977): branching temporal logic = logic of concurrency. The language of verification and model checking. Turing award in CS.

Logic of Knowledge: *a core AI topic,* $K_A(\varphi)$ = *"agent* A *knows* φ*", multiple modalities.*

Joe Halpern (1990s): Common knowledge operator cannot be expressed via individual knowledge operators. Problem: build a logic of knowledge that distinguishes hard and easy problems. Prime factorization example.

Basic systems of modal logic

System **K**:

A1. Propositional axioms and rules

A2. $\Box(F \Rightarrow G) \Rightarrow (\Box F \Rightarrow \Box G)$ *(distribution)*

Nec. Necessitation rule: $\dfrac{\vdash F}{\vdash \Box F}$

System **K4** is **K** $+$

A3. $\Box F \Rightarrow \Box\Box F$ *(positive introspection/transitivity)*

System **S4** is **K4** $+$

A4. $\Box F \Rightarrow F$ *(reflexivity)*

System **S5** is **S4** $+$

A5. $\neg\Box F \Rightarrow \Box(\neg\Box F)$ *(negative introspection)*

Some of derivations in **K** (hence in all other modal logics).

Theorem: $\Box$ *and* $\wedge$ *commute*

$A \Rightarrow (B \Rightarrow A \wedge B)$
$\Box(A \Rightarrow (B \Rightarrow A \wedge B))$
$\Box A \Rightarrow \Box(B \Rightarrow A \wedge B)$
$\Box A \Rightarrow (\Box B \Rightarrow \Box(A \wedge B))$
$(\Box A \wedge \Box B) \Rightarrow \Box(A \wedge B)$

$A \wedge B \Rightarrow A$
$\Box(A \wedge B \Rightarrow A)$
$\Box(A \wedge B) \Rightarrow \Box A$
$\Box(A \wedge B) \Rightarrow \Box B$
$\Box(A \wedge B) \Rightarrow (\Box A \wedge \Box B)$

Theorem: $\Box$ *factors out through* $\vee$:

$A \Rightarrow A \vee B$
$\Box(A \Rightarrow A \vee B)$
$\Box A \Rightarrow \Box(A \vee B)$
$\Box B \Rightarrow \Box(A \vee B)$
$(\Box A \vee \Box B) \Rightarrow \Box(A \vee B)$

But not $\Box(A \vee B) \Rightarrow (\Box A \vee \Box B)$!
Consider B to be $\neg A$. Whatever intended reading of modality you take $\Box(A \vee \neg A) \Rightarrow (\Box A \vee \Box \neg A)$ cannot be possibly true.

Modality dual to $\Box$: $\Diamond F \equiv \neg\Box\neg F$.

Intended semantics is derivative from the one for $\Box F$:

if $\Box F$ denotes "F holds in all possible situations",
then $\Diamond F$ stands for "F holds in at least one possible situation"

(the latter has been usually described as $\Box F$ denotes "F is necessary" and $\Diamond F$ stands for "F is possible")

Exercise: **S4** $\vdash A \Rightarrow \Diamond A$ (thus **S4** $\vdash \Box A \Rightarrow \Diamond A$).

Indeed: **S4** $\vdash \Box\neg A \Rightarrow \neg A$, **S4** $\vdash \neg\neg A \Rightarrow \neg\Box\neg A$, **S4** $\vdash A \Rightarrow \neg\Box\neg A$.

In many respects modal logics behave like normal logical systems. In particular, they are closed under substitution:

If $\Gamma(p) \vdash F(p)$ *then* $\Gamma(p/A) \vdash F(p/A)$ *for any* A

Modal logics admit equivalent substitution:
For **L=K**, **K4**, **S4**, **S5**, *if* $\mathbf{L} \vdash A \Leftrightarrow B$ *then* $\mathbf{L} \vdash F(p/A) \Leftrightarrow F(p/B)$
for any formula $F(p)$

NOTE: Deduction Theorem fails for **L=K**, **K4**, **S4**, **S5**. Indeed, in all of those logics $A \vdash \Box A$, by Necessitation, however, none of them derives $A \Rightarrow \Box A$. To prove that we need to develop some sort of negative test for **L**, for example, some sort of formal semantics true/false in a certain class of models along with a corresponding *soundness theorem*. Then by showing that F is false we can establish that F is not derivable.

One more example:

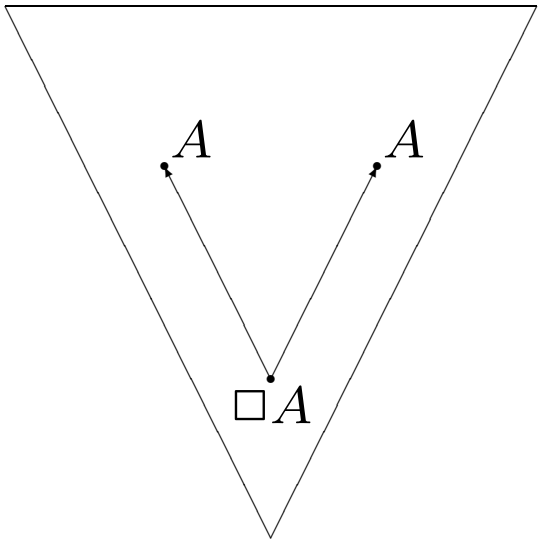
Derivation in **K4**, **S4**, **S5** that $F \Rightarrow \Box F$ holds not only for $F \equiv \Box A$ (transitivity axiom), but for $F \equiv \Box A \vee \Box B$ as well.

$\Box A \Rightarrow \Box A \vee \Box B$
$\Box(\Box A \Rightarrow \Box A \vee \Box B)$
$\Box A \Rightarrow \Box\Box A$
$\Box\Box A \Rightarrow \Box(\Box A \vee \Box B)$
$\Box A \Rightarrow \Box(\Box \vee \Box B)$
$\Box B \Rightarrow \Box A \vee \Box B$
$\Box(\Box B \Rightarrow \Box A \vee \Box B)$
$\Box B \Rightarrow \Box\Box B$
$\Box\Box B \Rightarrow \Box(\Box A \vee \Box B)$
$\Box B \Rightarrow \Box(\Box A \vee \Box B)$
$\Box A \vee \Box B \Rightarrow \Box(\Box A \vee \Box B)$

Possible Worlds Semantics by Saul Kripke.

Classical logic, propositional and quantified alike, gives a static picture of the world. A classical interpretation (model) is an assignment of truth values to atoms of the language. Modal logic has a striking ability to capture adequately a very natural semantics of "possible worlds" which can be traced back to Leibniz. The possible worlds universe consists of a collection of classical models W connected by a binary accessibility relation $R(a, b)$ "world b is accessible from world a". In other worlds, the possible worlds constitute an ordered graph, not necessarily finite. Whereas classical connectives operate within individual worlds (i.e. nodes in W), modality reaches out to all the worlds accessible from a given one (possible worlds):

$\Box F$ *holds in* a *iff* F *holds in all* b*'s accessible from* a.



Model Kripke is a triple $K = (W, R, \models)$, where W is a nonempty set (elements of which are called "possible worlds"), R a binary relation on W, and $\models$ a truth assignment having form: "world$\models$formula" such that each propositional letter gets some truth value in any world from W. We assume also that for any $x \in W$ both $x \models \texttt{true}$ and $x \not\models \texttt{false}$.

The definition of $x \models F$ (read as *a formula F is true in a world x*, or x *forces* F) goes by induction on F:

$x \models A \wedge B$ iff "$x \models A$ and $x \models B$"

$x \models A \vee B$ iff "$x \models A$ or $x \models B$"

$x \models \neg A$ iff "$x \not\models A$"

$x \models \Box A$ iff "$y \models A$ for all y such that $R(x, y)$"

By default, we assume that $A \Rightarrow B$ stands for $\neg A \vee B$, thus imposing the classical truth tables on boolean connectives at every given node. From the definition it is clear that a Kripke model is a collection of classical models connected by some sort of binary "accessibility" relation.

Modality $\Box$ is the only connective able to reach out to other possible worlds, i.e. nodes of the model accessible from a given one.

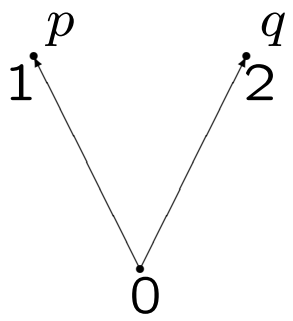
We may regard $\Box F$ as a sort of restricted universal quantifier "for all possible worlds F holds". It turns out that such limited quantification enables us to express some important features like time and process termination without compromising the decidability of the propositional logic.

$\Diamond F$ holds in x iff F holds in some y accessible from x.

Example

Consider a three-element "V-shaped" model with $W = \{0, 1, 2\}$ given by an oriented graph below. According to this graph, $R(0,1)$, $R(0,2)$, and neither of $R(1,2)$, $R(2,1)$, $R(1,0)$, $R(2,0)$, $R(0,0)$, $R(1,1)$, $R(2,2)$ holds.

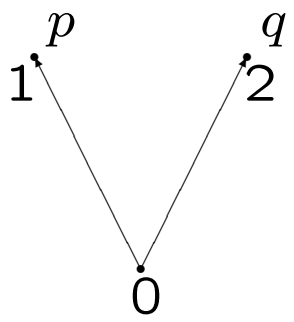
Notational convention: we label the nodes with propositional variables true at a given node. By default, all variables not listed next to a node are assumed false at this node. In particular, $1 \models p$, $2 \models q$, $1 \not\models q$, $2 \not\models p$, $0 \not\models p$, $0 \not\models q$, and all other variables are false at all nodes.



Question: for each of the formulas $\Box p$, $\Box q$, $\Box(p \wedge q)$, $\Box p \wedge \Box q$, $\Box(p \vee q)$, $\Box p \vee \Box q$, list the nodes where this formula is true.

Answer:

$\Box p$ is true at 1 and 2, but not at 0. Indeed, the set of accessible worlds for either 1 or 2 is empty, thus *FOR ALL* worlds accessible from each of them p holds. $\Box p$ is false at 0, since p fails at 2 which is accessible from 0.



Likewise, $\Box q$ holds at 1 and 2, but not in 0.

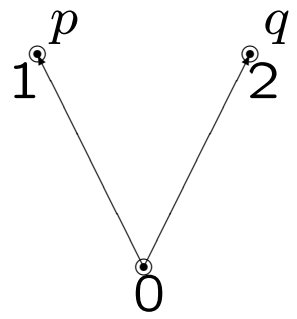
Formula $p \wedge q$ is false at every node. Formula $\Box(p \wedge q)$ is true at 1 and 2, but not at 0, so do $\Box p \wedge \Box q$ and $\Box p \vee \Box q$.

Formula $p \vee q$ is true at 1 and 2, but not at 0. Formula $\Box(p \vee q)$ is true at every node. Indeed, it is true at 1 and 2 by trivial reasons (above), hence it also true at 0, since $p \vee q$ is true at every possible world for it.

Note, that $0 \not\models \Box(p \vee q) \Rightarrow (\Box p \vee \Box q)$!. Hence we have found a model where this formula fails.

Truth value of a modal formula very much depends upon specific details of accessibility relation.

For example, consider the same model as above, but with all nodes made *reflexive*, i.e. $R(0,0)$, $R(1,1)$, and $R(2,2)$ (we denote reflexive worlds by "circled" nodes, as on the picture). The same formulas now have quite a different meaning.



In particular, $\Box p$ is true at 1, but not at 0 and 2. Likewise, $\Box p$ is true at 2, but not at 0 and 1.

It turned out that each of the modal logics under consideration is complete with respect to a corresponding class of Kripke models which can be characterized by the property of accessibility relation only.

Definition. A formula F is true in a model K (notation: $K \models F$) if F holds at every node of K. A formula F is valid (in a given class of models) if it is true in every model (of this class).

Consolidated Soundness Theorem

- If **K** $\vdash F$ then F is valid in all models.
- If **K4** $\vdash F$ then F is valid in all transitive models.
- If **S4** $\vdash F$ then F is valid in all transitive reflexive models.
- If **S5** $\vdash F$ then F is valid in all transitive reflexive symmetric models .

Proof. A pretty straightforward induction on the length of derivation in a given logic. We first prove that axioms are true in every model. Then we check that rules when applied to formulas true in all models (of a given class) produce a formula true in every such model as well.

Soundness of **K**.

A1. Propositional axioms

are true at every node since each node is a classical model.

A2. $\Box(F \Rightarrow G) \Rightarrow (\Box F \Rightarrow \Box G)$ *(distribution)*

We have to prove that A2 is true at every node x of every model. Suppose $x \models \Box(F \Rightarrow G)$ and $x \models \Box F$, then for every y accessible from x both $F \Rightarrow G$ and F hold, hence G does. Since G holds for every y accessible from x, the formula $\Box G$ holds at x.

Modus Ponens: $\dfrac{F \Rightarrow G, \quad F}{G}$. Obviously holds at each node.

Nec.: $\dfrac{\vdash F}{\vdash \Box F}$ By contrapositive, suppose there is a model K, where $\Box F$ is false at some node x. Then there should be a node y (accessible from x), where F is false. Therefore, F is false in K.

Soundness of **K4**

A3. $\Box F \Rightarrow \Box\Box F$ *(positive introspection/transitivity)*

Suppose $x \models \Box F$. In order to establish that $x \models \Box\Box F$ consider any y accessible from x and check that $y \models \Box F$. To do this, we have to consider any z accessible from y and prove that $z \models F$. The latter holds since z is also accessible from x (transitivity!), and thus $x \models \Box F$ yields $z \models F$.

Soundness of **S4**

A4. $\Box F \Rightarrow F$ *(reflexivity)*

Suppose $x \models \Box F$. Then $y \models F$ for all y accessible from x, in particular, for $y = x$. Thus $x \models F$.

Soundness of **S5**

A5. $\neg\Box F \Rightarrow \Box(\neg\Box F)$ *(negative introspection)*

Suppose $x \models \neg\Box F$, then $y \not\models F$ for some y accessible from x. In order to establish that $x \models \Box\neg\Box F$ consider any z accessible from x and check that $z \models \neg\Box F$. Since accessibility here is symmetric, x is accessible from z. By transitivity, y is also accessible from z. Thus we have found a node y accessible from z and such that $y \not\models F$. Thus $z \models \neg\Box F$.

To show that $p \Rightarrow \Box p$ is not derivable in modal logic, it now suffices to build a countermodel $K = (W, R, \models)$ for this formula. Consider $W = \{0, 1\}$ and let accessibility be a complete graph on W, i.e. $R(0,0)$, $R(0,1)$, $R(1,0)$, $R(1,1)$. Put $0 \models p$ and $1 \not\models p$. Clearly, K is a legitimate **S5** model, since R is an equivalence relation on W.

Moreover, $0 \models p$, but $0 \not\models \Box p$, since $1 \not\models p$ and 1 is accessible from 0. Therefore, $0 \not\models p \Rightarrow \Box p$.

By the soundness theorem, **S5** $\nvdash p \Rightarrow \Box p$, thus none of the other logics **K**, **K4**, **S4** does.

Consolidated Completeness Theorem

- **K** $\vdash F$ iff F is valid in all models.
- **K4** $\vdash F$ iff F is valid in all transitive models.
- **S4** $\vdash F$ iff F is valid in all transitive reflexive models.
- **S5** $\vdash F$ iff F is valid in all transitive reflexive symmetric models .

Proof. By the maximal consistent sets construction (sometimes called *canonical model*. A bit too long for our course.

Exercise. Prove that all logics **K**, **K4**, **S4**, **S5** are distinct. Hint: show that each next axiom is not derivable in the previous system, use models.

Temporal Logic - one of the most important brands of modal logic. We begin with the logic of *linear time*.

Linear Timeline is the set of natural numbers $0, 1, 2, 3, \ldots$, representing e.g. sequential computational process. It may also be regarded as a Kripke model where $R(x, y)$ is $x \leq y$.

Propositional Linear Temporal Logic (**PLTL**) has the following temporal connectives: `F`, `G`, `X`, and `U`

`G`p is *always* p, nothing but a new name for $\Box p$

`F`p is *sometime* p, a dual to `G`p, i.e. an analogue of $\Diamond p$

`X`p is *nexttime* p. Holds at n iff p holds at $n + 1$

p`U`q is p *until* q. Holds at n iff q does eventually hold at some $m \geq n$ and p holds everywhere at n or later prior to q.

Dependencies, definable connectives:

$\mathtt{F}p$ abbreviates $(\mathtt{true}\ \mathtt{U}p)$ (Exercise!)

$\mathtt{G}p$ is dual to $\mathtt{F}p$

$\mathtt{F}^{\infty}p$ - *infinitely often* p, abbreviates $\mathtt{GF}p$

$\mathtt{G}^{\infty}p$ - *almost everywhere* p, abbreviates $\mathtt{FG}p$

Definition. **PLTL** is defined not as a deductive system, but rather as a set of tautologies in a given language. This logic is decidable, and can be axiomatized by a finite set of schemes.

Examples: $p \Rightarrow \mathtt{F}q$ and $\mathtt{G}(p \Rightarrow \mathtt{F}q)$ are satisfiable, but not valid.

$\mathtt{G}(p \Rightarrow \mathtt{F}q) \Rightarrow (p \Rightarrow \mathtt{F}q)$ is valid, but its converse is only satisfiable.

$p \wedge \mathtt{G}(p \Rightarrow \mathtt{X}p) \Rightarrow \mathtt{G}p$ is valid, and called a temporal formulation of mathematical induction.

Logics of branching time - model of concurrency

Timeline a tree like discrete structure. Path quantifiers
$\mathtt{A}$ - "for all future paths"; $\mathtt{E}$ - "for some future paths. Examples:
$\mathtt{AF}^{\infty}p$ = "along each future path p happens infinitely often".
$\mathtt{EG}^{\infty}p$ = "along some future path p happens almost everywhere".

Reactive systems: operating systems, network communication protocols, air traffic control systems, etc. Normal behavior - arbitrary long, possible nonterminating computations.

Fairness: *each process is executed infinitely often* - $\mathtt{F}^{\infty}$
Safety property: *nothing bad happens* - $\mathtt{G}$
Liveness property: *something good will happen* - $\mathtt{F}$, $\mathtt{F}^{\infty}$, and $\mathtt{G}^{\infty}$