Logic II
(LGIC 320 / MATH 571 / PHIL 412)
Lecture Notes by Stepan Kuznetsov
University of Pennsylvania, Spring 2017

Modal Logic

1. The Language and Kripke Semantics

For simplicity, we consider only one modality (polymodal systems are outside the scope of our course). Also we do only the propositional case.

Modal formulae are built from propositional variables ($p, q, r, \ldots$) using standard logical (Boolean) connectives, $\vee$, $\wedge$, $\to$, $\bot$, and one extra unary operation, $\Box$, which is the *modality.* The interpretation of $\Box A$ is quite diverse: it could mean "A is necessarily true", "A is provable", "I believe that A is true", "I know that A is true", "A will be true tomorrow" etc. Different meanings of the modalities suggest different principles to be taken as axioms. Therefore we consider many modal logics and try, if possible, to construct a united theory for them.

First we define a formal semantic framework for modal logic, namely, Kripke semantics.

Definition. A *Kripke frame* is a structure $\mathcal{F} = \langle W, R \rangle$, where W is a non-empty set of *possible worlds,* and $R \subseteq W \times W$ is a binary relation on W.

Note that, contrary to the intuitionistic situation, here R is *arbitrary,* not necessarily a preorder.

Definition. A *Kripke model* is a structure $\mathcal{M} = \langle \mathcal{F}, v \rangle$, where $\mathcal{F}$ is a Kripke frame and $v \colon W \times \mathrm{Var} \to \{0, 1\}$ is an arbitrary *valuation function.*

Again, we don't impose any monotonicity conditions on v.

Definition. The truth of a formula A in a world $x \in W$ is defined recursively:

$$\begin{aligned}
\mathcal{M}, x \Vdash p &\iff v(x,p) = 1\\
\mathcal{M}, x \Vdash A \vee B &\iff \mathcal{M}, x \Vdash A \text{ or } \mathcal{M}, x \Vdash B\\
\mathcal{M}, x \Vdash A \wedge B &\iff \mathcal{M}, x \Vdash A \text{ and } \mathcal{M}, x \Vdash B\\
\mathcal{M}, x \Vdash A \to B &\iff \mathcal{M}, x \nVdash A \text{ or } \mathcal{M}, x \Vdash B\\
\mathcal{M}, x \nVdash \bot &\\
\mathcal{M}, x \Vdash \Box A &\iff \mathcal{M}, y \Vdash A \text{ for all } y \in R(x)
\end{aligned}$$

We also introduce three abbreviations: $\neg A \equiv (A \to \bot)$, $\top = \neg\bot$, and $\Diamond A \equiv \neg\Box\neg A$.

$$\mathcal{M}, x \Vdash \Diamond A \iff \mathcal{M}, y \Vdash A \text{ for some } y \in R(x).$$

$\Diamond$ is the dual modality to $\Box$.

Note that the following formula,

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

called the *normality* axiom, is true in all Kripke models.

Definition. The logic of frame $\mathcal{F}$, denoted by $\mathrm{Log}(\mathcal{F})$, is the set of all modal formulae that are true in all worlds of all models on the frame $\mathcal{F}$. The logic for a class of frames, $\mathcal{C}$, is

$$\mathrm{Log}(\mathcal{C}) = \bigcap_{\mathcal{F}\in\mathcal{C}} \mathrm{Log}(\mathcal{F}).$$

Definition. A *normal modal logic* is a set of formulae that includes all propositional tautologies, the normality axiom, and is closed under substitution and the following rules:

$$\frac{A \quad A \to B}{B}\ \mathrm{MP} \qquad\qquad \frac{A}{\Box A}\ \mathrm{Nec}$$

$\mathrm{Log}(\mathcal{C})$ is always a normal modal logic.

Definition. A normal modal logic $\mathcal{L}$ is *Kripke complete,* if $\mathcal{L} = \mathrm{Log}(\mathcal{C})$ for some $\mathcal{C}$.

Definition. For a normal modal logic $\mathcal{L}$ let $\mathrm{Mod}(\mathcal{L}) = \{\mathcal{F} \mid \mathcal{F} \Vdash A \text{ for every } A \in \mathcal{L}\}$.

The mappings Log and Mod are *contravariant:* if $\mathcal{C}_1 \subseteq \mathcal{C}_2$, then $\mathrm{Log}(\mathcal{C}_1) \supseteq \mathrm{Log}(\mathcal{C}_2)$, and if $\mathcal{L}_1 \subseteq \mathcal{L}_2$, then $\mathrm{Mod}(\mathcal{L}_1) \supseteq \mathrm{Mod}(\mathcal{L}_2)$.

Proposition 1. *$\mathcal{L}$ is Kripke complete iff $\mathcal{L} = \mathrm{Log}(\mathrm{Mod}(\mathcal{L}))$.*

Proof. The "if" part is by definition. For the "only if" part note that always $\mathcal{L} \subseteq \mathrm{Log}(\mathrm{Mod}(\mathcal{L}))$ (since $\mathcal{L}$ is true in all frames from $\mathrm{Mod}(\mathcal{L})$). On the other hand, $\mathcal{L} = \mathrm{Log}(\mathcal{C})$ for some $\mathcal{C}$, and $\mathcal{C} \subseteq \mathrm{Mod}(\mathcal{L})$ (since all formulae of $\mathcal{L}$ are true in any frame from $\mathcal{C}$). Therefore, $\mathcal{L} = \mathrm{Log}(\mathcal{C}) \supseteq \mathrm{Log}(\mathrm{Mod}(\mathcal{L}))$. $\Box$

Thus, a Kripke complete modal logic is a normal modal logic that is complete w.r.t. the class of its frames.

The minimal normal modal logic is denoted by **K**. Since it is true in all Kripke models, $\mathrm{Mod}(\mathbf{K}) = \{\text{all Kripke frames}\}$. Later we'll show that **K** (among other logics) is Kripke complete, i.e. $\mathrm{Log}(\text{all Kripke frames}) = \mathbf{K}$.

2. Classic Examples

In this section we briefly describe several well-known modal axioms (principles) and the corresponding classes of Kripke frames (the proof is left as an exercise).

reflexivity	$\Box p \to p$	all frames with reflexive R
transitivity	$\Box p \to \Box\Box p$	all frames with transitive R
symmetry	$\Diamond\Box p \to p$	all frames with symmetric R
seriality	$\Diamond\top$	all frames such that $R(x)$ is non-empty for every x

By $\mathcal{L}_1 + \mathcal{L}_2$ we denote the closure of $\mathcal{L}_1 \cup \mathcal{L}_2$ under the MP and Nec inference rules. It is easy to see that $\mathrm{Mod}(\mathcal{L}_1 + \mathcal{L}_2) = \mathrm{Mod}(\mathcal{L}_1) \cap \mathrm{Mod}(\mathcal{L}_2)$. For example, the class of frames for $\mathbf{K} + \Diamond\Box p \to p + \Box p \to p$ is the class of all symmetric reflexive frames. Some important combinations have specific names:

$$
\begin{aligned}
\mathbf{T} &= \mathbf{K} + \text{reflexivity} \\
\mathbf{K4} &= \mathbf{K} + \text{transitivity} \\
\mathbf{S4} &= \mathbf{K} + \text{reflexivity} + \text{transitivity} \\
\mathbf{K5} &= \mathbf{K} + \text{transitivity} + \text{symmetry} \\
\mathbf{S5} &= \mathbf{K} + \text{reflexivity} + \text{transitivity} + \text{symmetry} \\
\mathbf{D} &= \mathbf{K} + \text{seriality} \\
& \dots
\end{aligned}
$$

In the next section we show completeness for all these modal logics using one technique.

3. Canonical Models and Canonicity

The completeness proof strategy is basically the same as for intuitionistic logic.

Definition. Let $\mathcal{L}$ be a normal modal logic. A *theory* over $\mathcal{L}$ is a set of formulae that includes $\mathcal{L}$ is closed under MP (but maybe not Nec). A theory is *consistent,* if it doesn't include $\bot$. A theory is *complete,* if for any formula A it includes A or $\neg A$.

The standard saturation lemma holds: any consistent set of formulae that includes $\mathcal{L}$ can be embedded into a complete consistent theory.

Definition. The *canonical frame* for $\mathcal{L}$ is $\mathcal{F}_{\mathcal{L}} = \langle W_{\mathcal{L}}, R_{\mathcal{L}} \rangle$, where $W_{\mathcal{L}}$ is the set of all complete consistent theories over $\mathcal{L}$, and $xR_{\mathcal{L}}y$ if for every $\Box A \in x$ we have $A \in y$.

Definition. The *canonical model* for $\mathcal{L}$ is a model $\mathcal{M}_{\mathcal{L}}$ on the frame $\mathcal{F}_{\mathcal{L}}$ with the valuation function $v_{\mathcal{L}}$ such that $\mathcal{M}_{\mathcal{L}}, x \Vdash p \iff p \in x$.

Next, we prove the *main semantic lemma:*

Lemma 2. $\mathcal{M}_{\mathcal{L}}, x \Vdash A \iff A \in x.$

Proof. The only interesting case is $A = \Box B$. Then if $\Box B \in x$, for any $y \in R_{\mathcal{L}}(x)$ we have $B \in y$, and therefore by induction $y \Vdash B$. Hence $x \Vdash \Box B$.

For the reverse direction, suppose $\Box B \notin x$. Take the set $\mathcal{L} \cup \{C \mid \Box C \in x\} \cup \{\neg B\}$. This set is consistent: otherwise $\mathcal{L} \vdash C_1 \to (C_2 \to \dots \to (C_k \to B)\dots)$, by Nec $\mathcal{L} \vdash \Box(C_1 \to (C_2 \to \dots \to (C_k \to B)\dots))$, and by normality $\mathcal{L} \vdash \Box C_1 \to (\Box C_2 \to \dots \to (\Box C_k \to \Box B)\dots)$, and $\Box B \in x$. Saturate this set and obtain y. By definition, $xR_{\mathcal{L}}y$ and $\neg B \in y$. Therefore, $y \Vdash \neg B$, whence $y \nVdash B$ and $x \nVdash \Box B$. ☐

Then, if $\mathcal{L} \nvdash A$, then $\mathcal{L} \cup \{\neg A\}$ can be saturated to a world $x \in W_{\mathcal{L}}$, and therefore $\mathcal{M}_{\mathcal{L}} \nVdash A$. It looks like this shows Kripke completeness for an arbitrary $\mathcal{L}$, but things are not that simple. The problem is that in the definition of completeness we use *frames* rather than concrete *models.* It could be possible (and we present examples in the next section), that on the frame $\mathcal{F}_{\mathcal{L}}$ there exists another valuation v' such that $\langle \mathcal{F}_{\mathcal{L}}, v' \rangle \nVdash \mathcal{L}$, and therefore $\mathcal{F}_{\mathcal{L}} \notin \mathrm{Mod}(\mathcal{L})$.

Definition. $\mathcal{L}$ is *canonical,* if it is true in its canonical frame, $\mathcal{F}_{\mathcal{L}}$.

Since for a canonical logic $\mathcal{L} = \mathrm{Log}(\{\mathcal{F}_{\mathcal{L}}\})$, **every canonical logic is complete.**

Proposition 3. *If $\mathcal{L}_1$ and $\mathcal{L}_2$ are canonical, then so is $\mathcal{L}_1 + \mathcal{L}_2$.*

Proof. First we show that if $\mathcal{L} \subseteq \mathcal{L}'$, then $\mathcal{F}_{\mathcal{L}'}$ is a *conic subframe* of $\mathcal{F}_{\mathcal{L}}$, i.e., $W_{\mathcal{L}'} \subseteq W_{\mathcal{L}}$, the relation is induced, and $W_{\mathcal{L}'}$ is closed under $R_{\mathcal{L}}$. This is done straightforwardly by definition. Therefore, $\mathrm{Log}(\mathcal{F}_{\mathcal{L}'}) \supseteq \mathrm{Log}(\mathcal{F}_{\mathcal{L}})$ (truth is the same, the set of worlds gets smaller).

Now we need to show that $\mathcal{L}_1$ and $\mathcal{L}_2$ are true in $\mathcal{F}_{\mathcal{L}_1+\mathcal{L}_2}$. We know that $\mathcal{L}_i$ is true in $\mathcal{F}_{\mathcal{L}_i}$, and therefore it is true in its conic subframe $\mathcal{F}_{\mathcal{L}_1+\mathcal{L}_2}$ $\square$

This proposition, though trivial, is important: it allows to show canonicity independently for different axioms. For example, if we show that the transitivity, seriality, reflexivity, and symmetry axioms are canonical (exercise!), we prove that all the logics listed in the previous section are canonical and therefore Kripke complete.

4. The Gödel – Löb Logic

In this section we present an interesting logic, which is, by the way, an example of a non-canonical (but yet Kripke complete) normal modal logic. This is the system

$$\mathbf{GL} = \mathbf{K} + \Box(\Box p \to p) \to \Box p.$$

Lemma 4. $\mathbf{GL} \vdash$ *transitivity.*

Proof. Exercise. $\square$

Lemma 5. *The class* $\mathrm{Mod}(\mathbf{GL})$ *is the class of all transitive Kripke frames without infinite chains of the form $x_1 R x_2 R x_3 R \ldots$ (in particular, without cycles and reflexive points).*

Proof. If $\mathcal{M}, x \nVdash \mathbf{GL}$, then $x \Vdash \Box(\Box p \to p)$ and $x \nVdash \Box p$. The latter means that there exists a world $x_1 \in R(x)$ such that $x_1 \nVdash p$. On the other hand, $x_1 \Vdash \Box p \to p$, and therefore $x_1 \nVdash \Box p$. Moreover, by transitivity $x \Vdash \Box\Box(\Box p \to p)$, and therefore $x_1 \Vdash \Box(\Box p \to p)$. Now x_1 is in the same situation as x and we can continue *ad infinitum:* x_2, x_3, ... Thus, if $\mathbf{GL}$ is not true in a model, its frame has an infinite chain or is not transitive.

For the vice versa direction, we use the previous lemma to establish transitivity. Then, if the frame has an infinite chain, let p be false along this chain and true elsewhere. Then $\Box p$ is also false in all worlds of the chain, therefore $\Box p \to p$ is true everywhere, and in the beginning of the chain we have $\Box(\Box p \to p)$, but not $\Box p$. $\square$

In order to prove that $\mathbf{GL}$ is not canonical, we first introduce another, more natural interpretation of $\mathbf{GL}$. Consider a Gödelian arithmetical theory T, correct w.r.t. standard model (for example, $T = \mathbf{PA}$), and let $\mathrm{Pr}_T(x)$ be Gödel's provability predicate for T. Then we interpret propositional variables of the modal language as arbitrary closed arithmetical formulae, commute with Boolean operations, and let $(\Box A)^* = \mathrm{Pr}_T(\ulcorner A^* \urcorner)$, where $\ulcorner \varphi \urcorner$ is the Gödel number of φ. The Hilbert – Bernays conditions on Pr and Löb's theorem (which is the formalised version of the 2-nd Gödel theorem for the theory $T \cup \{\neg\varphi\}$) show *correctness* of $\mathbf{GL}$ w.r.t. this interpretation:

$$\text{if } \mathbf{GL} \vdash A, \text{ then } T \vdash A^*.$$

Note that reflexivity is not provable in $\mathbf{GL}$: although $\mathrm{Pr}_T(\ulcorner \varphi \urcorner)$ entains φ, if T is a correct theory, this fact is not provable in T unless φ is provable itself (Löb). However, this strange status of reflexivity leads to reflexive points in the canonical frame $\mathcal{F}_{\mathbf{GL}}$. Indeed, let $\mathbb{N}$ be the standard interpretation of arithmetic, and let $x_0 = \{A \mid \mathbb{N} \vDash A^*\}$. This x_0 is a complete theory over $\mathbb{GL}$ (a world of $\mathcal{F}_{\mathbf{GL}}$) and, moreover, if $\Box A \in x$, then $A \in x$ (because $\mathbb{N} \vDash T$). Therefore, x_0 is a reflexive point, $x_0 R x_0 R x_0 R \dots$ is an infinite chain, whence $\mathcal{F}_{\mathbf{GL}}$ doesn't satisfy $\mathbb{GL}$, therefore this logic isn't canonical.