Propositional Modal Logic*

John MacFarlane—Philosophy 142

February 15, 2011

1 Grammar

We add two new one-place sentential connectives to the language of propositional logic: '$\Box$' and '$\Diamond$'. Grammatically these behave just like $\neg$: if ϕ is a formula, then $\ulcorner \Box\phi \urcorner$ and $\ulcorner \Diamond\phi \urcorner$ are formulas.

You can read '$\Box$' as "necessarily" and '$\Diamond$' as "possibly." But modal logics, like other formal systems, can have many applications. Depending on the application, they might have many different meanings, for example:

$\Box$	$\Diamond$
it is logically necessary that	it is logically possible that
it could not have failed to happen that	it might have happened that
it must be the case that	it might be the case that
it is now settled that	it is still possible that
it is obligatory that	it is permitted that
it is provable that	it is not refutable that
A believes that	

Sometimes the letters 'L' and 'M' are used instead of '$\Box$' and '$\Diamond$'. Sometimes an operator for contingency is defined: $\nabla\phi =_{def} (\Diamond\phi \wedge \Diamond\neg\phi)$.

2 Semantics

2.1 Models

Our models for classical propositional logic were just assignments of truth values to the propositional constants—what is represented by the rows of a truth table. Models for modal logic must be more complex. A *model* consists of a frame and a valuation. A *frame* is a set of *worlds*, with one among them designated the "actual world," and an *accessibility relation* defined on those worlds. A *valuation* is a function that maps a propositional constant and a world to a truth value. More precisely:

> A **model** for modal propositional logic is a quadruple $\langle W, R, @, V\rangle$, where W is a nonempty set of objects (the "worlds"), R is a relation defined on W, $@$ is a member of W, and V is a function that assigns a truth value to each pair of a propositional constant and a world. The triple consisting of W, R, and $@$ is sometimes called a **frame**.

*I am indebted in my presentation to G. E. Hughes and M. J. Cresswell, *A New Introduction to Modal Logic* (London: Routledge, 1996) and to to Rod Girle, *Modal Logics and Philosophy* (Montreal: McGill-Queen's, 2000).

In many standard applications of modal logic, you can think of the worlds as "possible worlds"—ways things could be. You can think of the valuation function as telling us which propositions would be true in which possible worlds, that is, which would be true if things were a certain way. The "actual world" @ represents the way things *actually* are (according to the model).

You can think of the accessibility relation as embodying a notion of relative possibility. The worlds that are "accessible" from a given world are those that are possible relative to it. If this is too abstract, you can join Hughes and Cresswel in thinking of the worlds as seats, and the accessibility relation as the relation that holds between two seats if someone sitting in the first can see the person sitting in the second.

There are a lot of controversies about how we should think of possible worlds, metaphysically speaking: whether we should think of them as concrete worlds or as abstract models or sets of sentences, for example. For the most part, we can ignore these controversies when we're just doing logic.

2.2 Truth in a model for modal formulas

We define truth in a model for modal formulas in terms of quantification over worlds. Possibility is understood as truth in some accessible world, and necessity as truth in all accessible worlds. Here '$\vDash^{w}_{\mathcal{M}} \phi$' means '$\phi$ is true in model $\mathcal{M}$ at world w'. All the action is in the last two clauses:

- If ϕ is a propositional constant, $\vDash^{w}_{\langle W,R,@,V\rangle} \phi$ iff $V(\phi, w) = \text{True}$.
- $\vDash^{w}_{\langle W,R,@,V\rangle} \ulcorner\neg\phi\urcorner$ iff $\nvDash^{w}_{\langle W,R,@,V\rangle} \phi$.
- $\vDash^{w}_{\langle W,R,@,V\rangle} \ulcorner\phi \wedge \psi\urcorner$ iff $\vDash^{w}_{\langle W,R,@,V\rangle} \phi$ and $\vDash^{w}_{\langle W,R,@,V\rangle} \psi$.
- $\vDash^{w}_{\langle W,R,@,V\rangle} \ulcorner\Diamond\phi\urcorner$ iff for some $w' \in W$ such that Rww', $\vDash^{w'}_{\langle W,R,@,V\rangle} \phi$.
- $\vDash^{w}_{\langle W,R,@,V\rangle} \ulcorner\Box\phi\urcorner$ iff for every $w' \in W$ such that Rww', $\vDash^{w'}_{\langle W,R,@,V\rangle} \phi$.

So far we have defined truth in a model at a world. We can define (plain) truth in a model in terms of this as follows:

A formula ϕ is true in a model $\langle W, R, @, V\rangle$ if $\vDash^{@}_{\langle W,R,@,V\rangle} \phi$.

That is, a formula is true at a model if it is true at the model's "actual world."

We can now define the logical properties as usual in terms of truth at a model. A sentence is *logically true* if it is true in all models; an argument is *valid* if the conclusion is true in every model in which all the premises are true; and so on.

2.3 The modal logic K

If we define the logical properties this way and make no further restrictions on what counts as a model, we get the modal logic **K**. **K** is the weakest of the modal logics we'll look at, and everything that is valid in **K** is valid in all the others.

Here are some formulas that are logically true in **K**:

(1) $\Box(P \wedge Q) \supset (\Box P \wedge \Box Q)$

(2) $(\Box P \wedge \Box Q) \supset \Box(P \wedge Q)$

(3) $\neg\Box P \equiv \Diamond\neg P$

(4) $\Box\neg P \equiv \neg\Diamond P$

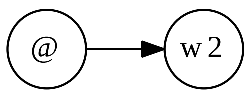
Can you see why they are true in all models? Think about (1) and (2) this way: if '$P \wedge Q$' is true in all accessible worlds, then it must be that P is true in all those worlds, and Q is true in all those worlds. The converse also holds: if P is true in all accessible worlds, and so is Q, then '$P \wedge Q$' is true in all accessible worlds.

Do you see the resemblance between (3) and (4) and the quantifier-negation equivalences? What explains this resemblance?

Here is a formula that is *not* logically true in **K**:

(5) $\Box P \supset P$

Can you see why not? Here is an invalidating model:



V	@	w_2
P	False	True

Here P is false in @, even though it is true at every world accessible from @.

Exercise:

2.3.1 Find a **K**-model in which '$\Box P \supset \Diamond P$' is false.

2.4 The modal logic D

If you add '$\Box\phi \supset \Diamond\phi$' as an axiom schema to **K**, you get a stronger logic **D**. (Stronger in the sense that it has more logical truths and more valid arguments.)

Since **D** is stronger than **K**, there must be **K**-models that are not **D**-models. (This makes it easier to find counterexamples in **K**.) In fact, **D**-models are **K**-models that meet an additional restriction: the accessibility relation must be *serial*. A relation R on W is *serial* iff $\forall w \in W \exists w' \in W Rww'$. What this means, intuitively, is that there are no "dead ends"—no worlds that can't "see" any worlds (including themselves).

With dead ends ruled out, '$\Box P \supset \Diamond P$' no longer has counterexamples.

Note that on the deontic interpretation of the modal operators, where '$\Box$' means "it is obligatory that" and '$\Diamond$' means "it is permissible that," '$\Box\phi \supset \Diamond\phi$' is essentially the principle "ought implies can." So **D** is a good logic for this interpretation. Note that in a deontic logic, we don't want '$\Box\phi \supset \phi$', since often what ought to be the case isn't the case.

2.5 The modal logic T

If you add '$\Box\phi \supset \phi$' as an axiom schema to **D**, you get a stronger logic **T**.

A **T**-model is a **K**-model whose accessibility relation is *reflexive*. A relation R on W is *reflexive* iff $\forall w \in W Rww$. That is, every world can see itself.

Since every reflexive accessibility relation is serial, every **T**-model is a **D**-model. The converse does not hold: there are **D**-models that are not **T**-models. Hence, every logical truth of **D** is a logical truth of **T**, but there are logical truths of **T** that are not logical truths of **D**.

Exercise:

2.5.1 Find a **T**-model in which '$\Box P \supset \Box\Box P$' is false.

2.5.2 Describe a **D**-model that is not a **T**-model.

2.6 The modal logic S4

If you add '$\Box\phi \supset \Box\Box\phi$' as an axiom schema to **T**, you get a stronger logic **S4**.

An **S4**-model is a **K**-model whose accessibility relation is *reflexive* and *transitive*. A relation R on W is *transitive* iff $\forall w_0, w_1, w_2 \in W((Rw_0w_1 \wedge Rw_1w_2) \supset Rw_0w_2)$.

Theorems of **S4** include '$\Diamond\Diamond P \supset \Diamond P$' and '$\Diamond\Box\Diamond P \supset \Diamond P$'.

Exercise:

2.6.1 Find an **S4**-model in which '$\Diamond P \supset \Box\Diamond P$' is false.

2.6.2 Find an **S4**-model in which '$\Diamond\Box P \supset P$' is false.

2.7 The modal logic B

If you add '$\phi \supset \Box\Diamond\phi$' as an axiom schema to **T**, you get a stronger logic **B**. (Note that neither **B** nor **S4** is stronger than the other; there are logical truths of **B** that are not logical truths of **S4**, and vice versa.)

A **B**-model is a **K**-model whose accessibility relation is *reflexive* and *symmetric*. A relation R on W is *symmetric* iff $\forall w, w' \in W(Rww' \equiv Rw'w)$.

Exercise:

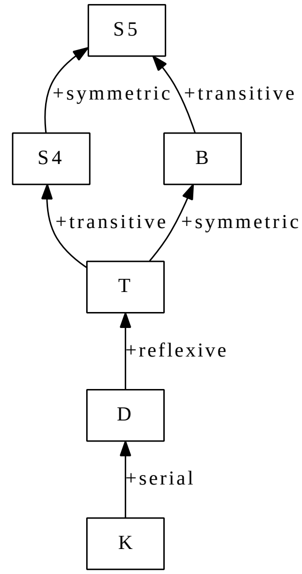
2.7.1 Find an **B**-model in which '$\Diamond P \supset \Box\Diamond P$' is false.

2.8 The modal logic S5

If you add '$\Diamond\phi \supset \Box\Diamond\phi$' as an axiom schema to **T**, you get a stronger logic **S5**. **S5** is stronger than both **S4** and **B**.

An **S5**-model is a **K**-model whose accessibility relation is *reflexive*, *symmetric*, and *transitive*. That is, it is an *equivalence relation*.

It is easy to see that if a formula can be falsified by an **S5**-model, it can be falsified by a *universal* **S5**-model—one in which every world is accessible from every other. So we get the same logic if we think of our models as just sets of possible worlds. Because the accessibility relation is an equivalence relation, we can more or less forget about it, and talk of necessity as truth in all possible worlds. **S5** is the most common modal logic used by philosophers.



Logic	Restrictions on accessibility relation	Characteristic axiom
K	—	—
D	serial	$\Box\phi \supset \Diamond\phi$
T	reflexive	$\Box\phi \supset \phi$
S4	reflexive, transitive	$\Box\phi \supset \Box\Box\phi$
B	reflexive, symmetric	$\phi \supset \Box\Diamond\phi$
S5	reflexive, symmetric, transitive	$\Diamond\phi \supset \Box\Diamond\phi$

Figure 1: Summary of the main systems of propositional modal logic.

Exercise:

2.8.1 Find an **S5**-model in which '$\Diamond P \supset \Box P$' is false.

2.9 Summary

The relationships between these six modal systems are summarized in Fig. 1.

3 Proofs

By supplementing our existing proof system for propositional logic with a few rules for the modal operators, we can get new proof systems for **T**, **S4**, and **S5**.

For all these systems, we'll need some modal-negation equivalences. These are replacement rules, so they can go in either direction and work even in subformulas (just like the parallel quantifier-negation equivalences):

Modal-negation equivalences (MNE)

$$\neg\Diamond\phi \iff \Box\neg\phi$$
$$\neg\Box\phi \iff \Diamond\neg\phi$$

We'll also need rules for $\Box$ Elim and $\Diamond$ Intr:

$$\frac{\Box\phi}{\phi}\ \Box\ \textbf{Elim} \qquad\qquad \frac{\phi}{\Diamond\phi}\ \Diamond\ \textbf{Intro}$$

Finally, we need some way of *introducing* a '$\Box$'. Obviously we can't have the converse of $\Box$ Elim, since that would let us prove every instance of $\ulcorner\phi \equiv \Box\phi\urcorner$, and our modal logic would be trivialized. We don't want to argue: "Fido is lying down; so, it is necessary that Fido be lying down."

The trick is to allow $\Box$ to be introduced only through a special kind of subproof. We will mark these subproofs with a small box to the left of the subproof line. (Some texts suggest putting a double line across the top of the subproof, or a box around the whole thing, to indicate that it is "sealed off" from the outside context, and you may want to do that.) If you have a modal subproof that ends with a formula ϕ, you can close off the subproof and write $\Box\phi$ on the next line, with justification "$\Box$ Intro."

$$\begin{array}{r|l|ll}
1 & \Box & \vdots & \\
2 & & \vdots & \\
3 & & \phi & \\
4 & \multicolumn{2}{l}{\Box\phi} & \Box \text{ Intro } 1\text{–}3
\end{array} \tag{6}$$

What is special about modal subproofs is that there are strict restrictions on the use of premises from outside the subproofs. In a subproof for conditional proof, it is just fine to appeal to premises from outside—for example,

$$\begin{array}{r|l|ll}
1 & \multicolumn{2}{l}{P} & \\ \hline
2 & & Q & \\ \cline{3-3}
3 & & P \wedge Q & \wedge \text{ Intro } 1, 2 \\
4 & \multicolumn{2}{l}{Q \supset (P \wedge Q)} & \supset \text{ Intro } 2\text{–}3
\end{array} \tag{7}$$

In our modal subproofs, by contrast, this won't be allowed:

$$\begin{array}{r|l|ll}
1 & \multicolumn{2}{l}{P} & \\ \hline
2 & \Box & \vdots & \\
3 & & Q & \\
4 & & P \wedge Q & \wedge \text{ Intro } 1, 3 \Leftarrow \textbf{ILLEGAL!} \\
5 & \multicolumn{2}{l}{\Box(P \wedge Q)} & \Box \text{ Intro } 2\text{–}4
\end{array} \tag{8}$$

Here's a way to think about the difference between regular and modal subproofs. Regular subproofs allow things to *enter* freely, but *exit* only according to strict rules. Modal subproofs impose restrictions both on entry *and* on exit.

The entry restrictions are given by the *modal reiteration rule.* These are the *only* rules that allow you to use premises outside the subproof. Other rules may be used only on premises within the modal subproof. The modal reiteration rule(s) available depend on the modal logic (in fact, this rule is the only one that changes between different systems):

Modal reiteration rules

$$\frac{\Box\phi}{\phi}\ \textbf{Modal reit T} \qquad \frac{\Box\phi}{\Box\phi}\ \textbf{Modal reit S4} \qquad \frac{\Diamond\phi}{\Diamond\phi}\ \textbf{Modal reit S5}$$

Note: in **S5**, you may use any of these rules. In **S4**, you may use the S4 or T rules. In **T**, you may only use the T rule.

A modal subproof may be started at any time, provided these entry restrictions are obeyed. There is no separate "hyp" or "flagging" step.

Here's an example of a proof in **T** of '$(\Box P \wedge \Box Q) \supset \Box(P \wedge Q)$':

$$\begin{array}{lll}
1 & \mid\ \Box P \wedge \Box Q & \\
 & \mid\overline{} & \\
2 & \mid\ \Box P & \text{Taut Con, 1} \\
3 & \mid\ \Box Q & \text{Taut Con, 1} \\
4 & \mid\ {}^{\Box}\!\mid\ P & \text{Modal Reit T, 2} \\
5 & \mid\ \ \ \mid\ Q & \text{Modal Reit T, 3} \\
6 & \mid\ \ \ \mid\ P \wedge Q & \text{Taut Con, 4, 5} \\
7 & \mid\ \Box(P \wedge Q) & \Box\ \text{Intro, 4–6} \\
8 & (\Box P \wedge \Box Q) \supset \Box(P \wedge Q) & \supset\ \text{Intro, 1–7}
\end{array} \tag{9}$$

Here's a proof in **S4** of '$(\Box P \vee \Box Q) \supset \Box(P \vee \Box Q)$':

$$\begin{array}{lll}
1 & \mid\ \Box P \vee \Box Q & \\
 & \mid\overline{} & \\
2 & \mid\ \mid\ \Box P & \\
 & \mid\ \mid\overline{} & \\
3 & \mid\ \mid\ {}^{\Box}\!\mid\ P & \text{Modal Reit T, 2} \\
4 & \mid\ \mid\ \ \ \mid\ P \vee \Box Q & \vee\ \text{Intro, 3} \\
5 & \mid\ \mid\ \Box(P \vee \Box Q) & \Box\ \text{Intro, 3–4} \\
6 & \mid\ \mid\ \Box Q & \\
 & \mid\ \mid\overline{} & \\
7 & \mid\ \mid\ {}^{\Box}\!\mid\ \Box Q & \text{Modal Reit S4, 6} \\
8 & \mid\ \mid\ \ \ \mid\ P \vee \Box Q & \vee\ \text{Intro, 7} \\
9 & \mid\ \mid\ \Box(P \vee \Box Q) & \Box\ \text{Intro, 7-8} \\
10 & \mid\ \Box(P \vee \Box Q) & \vee\ \text{Elim, 1, 2–5, 6–9} \\
11 & (\Box P \vee \Box Q) \supset \Box(P \vee \Box Q) & \supset\ \text{Intro, 1–10}
\end{array} \tag{10}$$

Exercises:

3.1 Use a deduction to show that the following argument is valid in **T**: $\Box(A \supset B), \Box(B \supset C), \Box(C \supset D), \neg\Diamond D, / \therefore \neg\Diamond A$.

3.2 Give deductions for the following in **S5**:

(a) $\Diamond\Diamond P \supset \Diamond P$

(b) $\Diamond(P \vee Q) \equiv (\Diamond P \vee \Diamond Q)$

3.3 For each of the following formulas, determine whether it is a logical truth of **T**, **S4**, and/or **S5**. Give countermodels when a formula is not a logical truth of a system, deductions when it is. (Check each formula against all three systems.)

(a) $\Box(P \supset \Box\Diamond P)$

(b) $\Box\Box P \vee \Box\neg\Box P$

(c) $\Diamond(P \vee Q) \supset \Diamond P$

(d) $\Diamond\Box P \supset \Box P$

(e) $\Diamond\Box\Diamond P \supset \Diamond P$

3.4 *Extra credit:* We have given you proof systems for **T**, **S4**, and **S5**. Can you come up with systems that make sense for **D** and **B**?