MODEL THEORETIC CONCEPTS IN MODAL LOGIC

Standard deontic logic (SDL) is a special case of propositional modal logic. We introduce it by presenting the latter. We rely on the textbook *Boxes and Diamonds* by Richard Zach (abbreviated *B&D*) which is freely available and part of the Open Logic Project. In the tutorial, we will work more closely on SDL.

1 Language

The language of standard deontic logic is generated by:

- atomic sentences (p, q, r, etc.) together with the propositional constant $\bot$
- Boolean connectives ($\&$, $\vee, \neg, \rightarrow, \leftrightarrow$)
- Unary deontic operators:

  $\bigcirc$ (for obligation)

  $\mathbf{P}$ (for permission)

  $\mathbf{F}$ (for prohibition)

Note 1: we use p, q, r for the atomic sentences of our language, $A, B, C, \ldots$ etc. as variables ranging over sentences; S, T as variables ranging over sets of sentences.

Note 2: the founding assumption of standard deontic logic is that $\bigcirc$ works like the $\Box$ of modal logic and $\mathbf{P}$ works like the $\Diamond$. (But if you are not familiar with this notation don't worry!).

2 Core Equivalences

These are standardly assumed equivalences between the three deontic operators. The table below expresses the row items in terms of the column items and negation.

	$\bigcirc A$	$\mathbf{P}A$	$\mathbf{F}A$
$\bigcirc A$		$\bigcirc A \leftrightarrow \neg\mathbf{P}\neg A$	$\bigcirc A \leftrightarrow \mathbf{F}\neg A$
$\mathbf{P}A$	$\mathbf{P}A \leftrightarrow \neg\bigcirc\neg A$		$\mathbf{P}A \leftrightarrow \neg\mathbf{F}A$
$\mathbf{F}A$	$\mathbf{F}A \leftrightarrow \bigcirc\neg A$	$\mathbf{F}A \leftrightarrow \neg\mathbf{P}A$	

We could use these equivalences to 'define operators away' or simply as design principles that any adequate deontic logic ought to deliver.

Exercise 1. Use $\bigcirc A \leftrightarrow \neg \mathbf{P} \neg A$ and $\mathbf{F} A \leftrightarrow \bigcirc \neg A$ to derive all the other equivalences.

3 Kripke Models for Modal Logic

Models for modal logic (and deontic logic in particular) are triples $\langle W, R, V \rangle$ with

- W a non-empty set of worlds
- R an accessibility relation over W
- V a valuation function (i.e. a function mapping atomic sentences of the language to sets of worlds).

In *deontic* logic specifically, the accessibility relation R is typically given an informal interpretation roughly like the following two.

Interpretation one: wRv iff v is ideal from the point of view of w

Interpretation two: wRv iff every (salient) norm that prevails in w is satisfied in v

In the tutorial, we will think about the particular shape that the accessibility relation must take for deontic logic. The central assumption is that it needs to be at least *serial*: that is, *every world must access some world*.

Exercise 2. Try to work informally (but abstractly) on these interpretations.

- what would it mean for a world w to be related to itself (i.e. wRw)?
- and what would it mean if world v was accessed by w (i.e. wRv) but did not access itself (i.e. it is not the case that vRv?)

4 Semantics for Modal Logic

The semantic module centers around an account of truth in a model $\mathcal{M}$ at a particular world w. When A is true in $\mathcal{M}$ at w, we write $\mathcal{M}, w \models A$.

Preliminary notes:

one: This definition is recursive (meaning that we start with the atomic sentences and we build up to more complex sentences).

two: We typically use the "compressed" notation $\mathcal{M}$ to refer to a model $\langle W, R, V \rangle$. When we want to talk about, say, the valuation function of $\mathcal{M}$ we write $V^{\mathcal{M}}$.

three: We only do *some* of the cases and leave the others as exercises.

Truth at a model-world pair. (cf. B&D, def 1.6)

Clause 0. Suppose A is $\bot$. Then $\mathcal{M}, w \not\models \bot$.

Clause 1. Suppose A is atomic. Then: $\mathcal{M}, w \models A$ iff $w \in V^{\mathcal{M}}(A)$

Clause 2. Suppose $A = B \wedge C$. Then $\mathcal{M}, w \models B \wedge C$ iff $\mathcal{M}, w \models B$ and $\mathcal{M}, w \models C$

Clause 3. Suppose $A = \neg B$. Then $\mathcal{M}, w \models \neg B$ iff $\mathcal{M}, w \not\models B$

Clause 4. Suppose $A = \bigcirc(B)$. Then $\mathcal{M}, w \models \bigcirc B$ iff for all v with wRv, $\mathcal{M}, v \models B$

Using the interdefinability of boolean connectives, we can uses clauses 2 and 3 to derive the clauses for $\vee$, $\rightarrow$ and $\leftrightarrow$

Using the interdefinability of $\bigcirc$ and **P** we can use clause 4 to derive the truth-conditions of permission claims. Recall that $\mathbf{P}A \leftrightarrow \neg\bigcirc\neg A$. So then you can reason:

	$\mathcal{M}, w \models \mathbf{P}A$	
iff	$\mathcal{M}, w \models \neg\bigcirc\neg A$	["*definition*"]
iff	$\mathcal{M}, w \not\models \bigcirc\neg A$	[*clause 3*]
iff	it's not the case that for all v with $wRv, \mathcal{M}, v \models \neg A$	[*unpacking clause 4*]
iff	there is a world v with wRv such that $M, v \models A$	[*basic logic in the metalanguage*]

Exercise 3. Derive, or in any case identify, the clauses for disjunction, conditional, biconditional and prohibition.

5 Other Semantic Concepts

The previous definition characterizes what it is to be true *in* a model *at* a world w. Other semantic concepts are also important in the project of characterizing modal validity.

5.1 Global Model Constraints

True everywhere. A is *true everywhere* in $\mathcal{M}$ iff for every world v in $W^{\mathcal{M}}$, $\mathcal{M}, v \models A$
True somewhere. A is *true somewhere* in $\mathcal{M}$ iff for some world v in $W^{\mathcal{M}}$, $\mathcal{M}, v \models A$

Where S is a set of sentences we can also say:

- S is true somewhere in $\mathcal{M}$ iff for some world v in $W^{\mathcal{M}}$, and for every sentence A in S, $\mathcal{M}, v \models A$

Important! For S to be true somewhere, all of S has to be true in the same world.

Exercise 4. Diagram a model with two worlds, making $p \vee q$ true everywhere while all of $p, \neg p, q, \neg q$ are true somewhere in the model.

5.2 Frames

Each model $\mathcal{M} = \langle W, R, V \rangle$ is associated with a *frame* $\mathcal{F}_{\mathcal{M}} = \langle W, R \rangle$.

Intuitively: the frame is the model "without" the valuation function.

Note: there is a many-one relation between models and frames (each model determines a frame, but there are many models corresponding to each frame).

models$_{\mathcal{F}}$ = the class of all models built on frame $\mathcal{F}$.
Valid on a frame. A sentence A is valid on a frame $\mathcal{F}$ iff for all models $\mathcal{M}$ in **models**$_{\mathcal{F}}$, A is true everywhere in $\mathcal{M}$.

Big, if slightly mysterian, idea. You'll make a big leap in modal logic if you start thinking of sentences as constraints on frames.

6 Standard model theoretic analyses of logical concepts

Satisfiability

A single sentence A (*/a set of sentences* S) is *satisfiable* in a class $\mathcal{C}$ of models iff there is a model $\mathcal{M}$ in $\mathcal{C}$ such that A (*/the entire set* S) is true somewhere in $\mathcal{M}$

Note: Satisfiability is a model theoretic analogue of consistency.

Validity (cf. B&D §1.7)

A single sentence A is *valid* in a class $\mathcal{C}$ of models iff for every model $\mathcal{M}$ in $\mathcal{C}$, A is true everywhere in $\mathcal{M}$

Exercise 5. B&D singles out two propositions about validity. One: if $\mathcal{C}' \subseteq \mathcal{C}$ and A is valid in $\mathcal{C}$, then $\mathcal{A}$ is valid in $\mathcal{C}'$. Two: if A is valid in $\mathcal{C}$, then $\Box A$ is valid in $\mathcal{C}$.

Entailment (cf. B&D §1.10, but modified)

An argument with premises in S and conclusion A is an *entailment* in a class $\mathcal{C}$ of models iff $S \cup \{\neg A\}$ is not satisfiable in $\mathcal{C}$.