Modal Logic

(Lecture Notes for Applied Logic)

Anton Hedin

October 2, 2008

Contents

1 Introduction

In (classical) propositional and predicate logic, every formula is either true or false in any model. But there are situations were we need to distinguish between different modes of truth, such as *necessarily true, known to be true, believed to be true* and *always true in the future* (with respect to time). For example, consider the sentence

"The math department is located on the fourth floor of the Ångström building".

It expresses something that is true today, but was false some years ago. Moreover, it might be false again some time in the future. Another example, whose mode of truth is more stable with respect to time, is

"The Earth has exactly one moon".

The sentence expresses something that is true and maybe will be true for ever in the future, but it is not necessarily true in the sense that there might as well have been two moons, or none for that matter (since we know this is possible for planets). However, most people would consider the statement

"The square root of 9 is 3"

as expressing something that is both necessarily true as well as always true. But it does not enjoy all modes of truth, for example it may not be believed to be true (for example by someone who is mistaken) or known to be true (for example by someone that hasn't learned mathematics).

There are also more practical examples were reasoning about different modes of truth is helpful. For example, think of a multi-agent system in computer science. There, each agent may have different knowledge about the system and even about other agents knowledge. In such a scenario a sentence is 'necessarily' true when known. It should be clear that not every sentence needs to be necessarily true in this sense.

In the examples given above we are using the same way of reasoning. A sentence φ, if true, will be so with respect to the current state of affairs, i.e.

how the world actually is, but (depending on φ) we might be able to conceive of a state of affairs (a different world) were φ is false, and if this is the case φ will not be necessarily true. These states of affairs can be points in time as in the first example, possible worlds as in the second example or states of knowledge of a person/agent as in the last two examples.

We will develop a general framework in which we will be able to reason about situations as the ones above. First we take a look at basic modal logic.

2 Basic Modal Logic

2.1 Syntax

The language of Basic Modal Logic is an extension of classical propositional logic. What we add are two unary connectives $\Box$ and $\Diamond$. We have a set *Atoms* of propositional letters $p, q, r, \ldots$, also called atomic formulas or atoms.

Definition 1. Formulas of basic modal logic are given by the following rule

$$\varphi ::= \bot \mid \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid (\varphi \to \varphi) \mid (\varphi \Leftrightarrow \varphi) \mid \Box\varphi \mid \Diamond\varphi.$$

where p is any atomic formula.

Examples of well formed formulas (wffs) are $(q \wedge \neg\Diamond p)$ and $(\Box p \to \Diamond\Box(r \vee \Box\top))$, while $p\Box\neg \to \Diamond p$ or $\vee p\Box q$ are clearly non-wffs!

Just as in predicate logic, the unary connectives bind most closely so that for example $\Box p \vee r$ is read as $(\Box p) \vee r$ and not $\Box(p \vee r)$.

The new connectives $\Box$ and $\Diamond$ are read 'box' and 'diamond' respectively, and are dual of each other similarly to how $\forall$ and $\exists$ are dual of each other in predicate logic (we will return to this later). And just as $\forall$ and $\exists$ are read as 'for every' and 'there exists' respectively, we will also want to give special readings for box and diamond. Although the readings will be different depending on the situation we want to study, i.e. what mode of truth we are interested in. For example in case we want to study necessity and possibility (as in the case of the second sentence above), $\Box$ is read as 'necessarily' and

$\Diamond$ as 'possibly'. In such a logic there are some formulas we might regard as being correct principles, for example $\Box\varphi \rightarrow \Diamond\varphi$ 'whatever is necessary is possible' or $\varphi \rightarrow \Diamond\varphi$ 'whatever is, is possible'. However, other formulas may be harder to decide. Should $\varphi \rightarrow \Box\Diamond\varphi$ 'whatever is, is necessarily possible' be regarded as a general truth about necessity and possibility? A precise semantics will bring clarity to questions like these.

Remark 2.1. We could just as well have defined a formula φ by the following (shorter) rule

$$\varphi ::= \bot \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \rightarrow \varphi) \mid \Box\varphi,$$

and then, as is usually done, define the rest of the connectives from the ones given. In this case, diamond would be defined as $\Diamond := \neg\Box\neg$.

2.2 Semantics

We now want to give some mathematical content to our suggestive discussion above. A model in propositional logic is simply a valuation function assigning truth values to the set of atoms, i.e. a function

$$\nu : \mathit{Atoms} \rightarrow \{\top, \bot\}.$$

As we've hinted at in the discussion, we now want to consider models in which an atom can have different truth values at different states. Therefore:

Definition 2. A *model*, $\mathcal{M}$, in basic modal logic is a triple (W, R, L), where

- W is a set of *states* or *worlds*,
- R is a relation $R \subseteq W \times W$,
- and L is a function $L : W \rightarrow \mathcal{P}(\mathit{Atoms})$, called the *labelling function*.

These models are called Kripke models after Saul Kripke who was the first to introduce them in the 1950s. Intuitively $w \in W$ is a possible world and R is an accessibility relation between worlds. That is, wRw' (which we will use as shorthand for $(w, w') \in R$) means that w' is accessible from w. This intuition will be made more precise in the next definition. But first some examples.

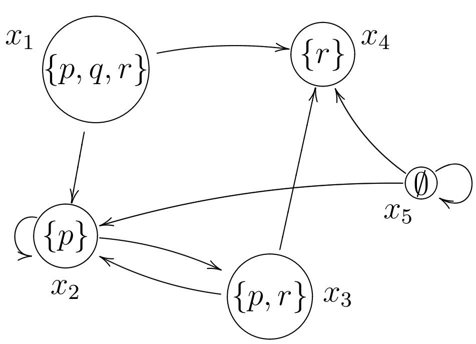
Example 1. Although the definition of a Kripke model might look somewhat complicated, we can use an easy graphical notation for such a model: Suppose $\mathcal{M} = (W, R, L)$ is a Kripke model where

$$\begin{aligned} W &= \{x_1, x_2, x_3, x_4, x_5\}, \\ R &= \{(x_1, x_2), (x_1, x_4), (x_2, x_2), (x_2, x_3), \\ &\qquad (x_3, x_2), (x_3, x_4), (x_5, x_5), (x_5, x_2), (x_5, x_4)\} \end{aligned}$$

and

$$L(x_i) = \begin{cases} \{p, q, r\}, & i = 1 \\ \{p\}, & i = 2 \\ \{p, r\}, & i = 3 \\ \{r\}, & i = 4 \\ \emptyset, & i = 5 \end{cases}$$

Then we can picture $\mathcal{M}$ as follows:



Where an arrow $x_i \to x_j$ means that $x_i R x_j$.

Example 2. A Kripke model $\mathcal{M} = (W, R, L)$ can be used to describe how truth values vary over time. A common example is when $W = \mathbb{N}$ and R is the ordinary ordering $\leq$ of the natural numbers.

$$0 \longrightarrow 1 \longrightarrow 2 \longrightarrow \cdots \longrightarrow n \longrightarrow \cdots$$

Then we can think of W as a set of points in time and R as the relation of being ahead in time. Then $L(t)$ will describe the truth values of propositions at time $t \in W$.

Example 3. If we consider the model $(\{*\}, \emptyset, L)$, we see that it is in fact just an ordinary model for propositional logic (when restricting our attention to formulas of propositional logic). Namely, we can define an ordinary valuation function $V : Atoms \to \{\top, \bot\}$ by

$$V(p) = \begin{cases} \top, & \text{if } p \in L(*) \\ \bot, & \text{if } p \notin L(*) \end{cases}$$

Now we will define what it means for a formula to be true at a state in a model.

Definition 3. Let $\mathcal{M} = (W, R, L)$ be a model in basic modal logic. Suppose $x \in W$ and φ is a formula. We will define when φ *is true in the world* x. This is done via a satisfaction relation $x \Vdash \varphi$ by structural induction on φ:

- $x \Vdash \top$
- $x \nVdash \bot$
- $x \Vdash p$ iff $p \in L(x)$
- $x \Vdash \neg\varphi$ iff $x \nVdash \varphi$
- $x \Vdash \varphi \wedge \psi$ iff $x \Vdash \varphi$ and $x \Vdash \psi$
- $x \Vdash \varphi \vee \psi$ iff $x \Vdash \varphi$ or $x \Vdash \psi$
- $x \Vdash \varphi \to \psi$ iff $x \Vdash \psi$ whenever $x \Vdash \varphi$
- $x \Vdash \varphi \leftrightarrow \psi$ iff $x \Vdash \varphi$ iff $x \Vdash \psi$
- $x \Vdash \Box\varphi$ iff for each $y \in W$ with xRy we have $y \Vdash \varphi$
- $x \Vdash \Diamond\varphi$ iff there exists $y \in W$ such that xRy and $y \Vdash \varphi$

When $x \Vdash \varphi$, we say that 'x satisfies/forces φ' or 'φ is true in world x/at state x'.

The first eight clauses are straightforward from propositional logic, the only difference being that an atom p can be true at many worlds x. The interesting cases are the ones for box and diamond. For $\Box\varphi$ to be true at x, φ must hold at every world y related to x, and for $\Diamond\varphi$ to be true at x there must be at least one world y related to x such that φ is true at y. Note that $\Box$ and $\Diamond$ act a bit like the quantifiers $\forall$ and $\exists$, but quantifiers over states instead of variables. The above interpretation of the logical constants of basic modal logic is usually called *possible worlds semantics*.

Example 4. Consider the model of example 1. According to definition 3 we have

$$x_2 \Vdash \Box p$$

since $x_2 R y$ implies that $y = x_2$ or $y = x_3$ and we have both $x_2 \Vdash p$ and $x_3 \Vdash p$, i.e. $p \in L(x_2)$ and $p \in L(x_3)$. Moreover we have

$$x_1 \models \Diamond(r \wedge \Box q)$$

since $x_1 R x_4$ and $x_4 \Vdash r$ and $x_4 \Vdash \Box q$ since there is no $y \in W$ such that $x_4 R y$.

Definition 4. A model $\mathcal{M} = (W, R, L)$ is said to satisfy a formula φ if every state in the model satisfies it. Thus, we write $\mathcal{M} \models \varphi$ if and only if $x \Vdash \varphi$ for every $x \in W$.

Example 5. Again considering the model $\mathcal{M}$ of example 1 we see that for example the modal formula $r \vee \Diamond p$ is satisfied by $\mathcal{M}$: $x_1, x_3, x_4 \Vdash r$ and $x_2, x_5 \Vdash \Diamond p$.

Next we define semantic entailment.

Definition 5. Let Γ be set of formulas. Then we say that Γ *semantically entails* a formula φ if for any world x in any model $\mathcal{M} = (W, R, L)$ we have $x \Vdash \varphi$ whenever $x \Vdash \psi$ for every $\psi \in \Gamma$. In that case we write $\Gamma \models \varphi$.

We will say that two formulas φ and ψ are semantically equivalent when they semantically entail each other, and then we write $\varphi \equiv \psi$.

Example 6. We have already seen that $\Box$ acts as a universal quantifier on worlds while $\Diamond$ acts as an existential quantifier on worlds. Therefore it may not be so surprising that we have the following semantical equivalences:

$$\Box\varphi \equiv \neg\Diamond\neg\varphi$$

$$\Diamond\varphi \equiv \neg\Box\neg\varphi$$

How can we see this? Well, let $\mathcal{M} = (W, R, L)$ be an arbitrary model and let $x \in W$ be any world. Suppose $x \Vdash \Box\varphi$, then $y \Vdash \varphi$ for every $y \in W$ such that xRy. So there cannot exist a world $y \in W$ such that xRy and $y \Vdash \neg\varphi$, but then $x \nVdash \Diamond\neg\varphi$. Hence $x \Vdash \neg\Diamond\neg\varphi$. Conversely, if $x \Vdash \neg\Diamond\neg\varphi$ then $x \nVdash \Diamond\neg\varphi$ so that there is no world y such that xRy and $y \Vdash \neg\varphi$. Hence, for any $y \in W$ such that xRy, we must have $y \Vdash \varphi$. But then $x \Vdash \Box\varphi$. Hence $\Box\varphi \equiv \neg\Diamond\neg\varphi$. The second equivalence follows easily from the first one.

Example 7. It is also not surprising that $\Box$ distributes over $\wedge$ and that $\Diamond$ distributes over $\vee$ but not the other way around. That is

$$\Box(\varphi \wedge \psi) \equiv (\Box\varphi \wedge \Box\psi)$$

$$\Diamond(\varphi \vee \psi) \equiv (\Diamond\varphi \vee \psi)$$

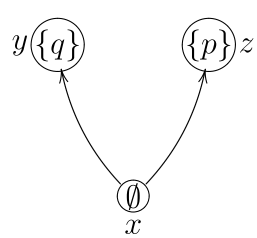
but

$$\Box((\varphi \vee \psi) \not\equiv (\Box\varphi \vee \Box\psi)$$

$$\Diamond(\varphi \wedge \psi) \not\equiv (\Diamond\varphi \wedge \Diamond\psi).$$

For the first equivalence, let $\mathcal{M} = (W, R, L)$ be an arbitrary model and let $x \in W$ be any world. Suppose $x \Vdash \Box(\varphi \wedge \psi)$, then for every y such that xRy, $y \Vdash \varphi$ and $y \Vdash \psi$. But then of course $x \Vdash \Box\varphi$ and $x \Vdash \Box\psi$, i.e. $x \Vdash \Box\varphi \wedge \Box\psi$. Likewise, if $x \Vdash \Box\varphi \wedge \Box\psi$, then for every y such that xRy, $y \Vdash \varphi$ and $y \Vdash \psi$, i.e. $y \Vdash \varphi \wedge \psi$ and hence $x \Vdash \Box(\varphi \wedge \psi)$.

To see that $\Box(\varphi \vee \psi) \not\equiv \Box\varphi \vee \Box\psi$, we consider the following kripke model:



We see that $x \Vdash \Box(p \vee q)$, since $y \Vdash p \vee q$ and $z \Vdash p \vee q$. However $x \nVdash \Box p \vee \Box q$ since $y \nVdash p$ and $z \nVdash q$.

Exercise 1. Show that $\Diamond(\varphi \vee \psi) \equiv (\Diamond\varphi \vee \psi)$ and that $\Diamond(\varphi \wedge \psi) \not\equiv (\Diamond\varphi \wedge \Diamond\psi)$.

Now, we need a notion of validity. In our case this will mean not just being true with respect to every valuation, but also with respect to every underlying relational structure (W, R). More precisely we have

Definition 6. We say that a formula φ is *valid* if it is true in every world of every model. We denote this by $\models \varphi$.

From the results in example 6 and 7 we have that the following formulas are valid

$$\Diamond\varphi \leftrightarrow \neg\Box\neg\varphi$$

$$\Box(\varphi \wedge \psi) \leftrightarrow (\Box\varphi \wedge \Box\psi)$$

$$\Diamond(\varphi \vee \psi) \leftrightarrow \Diamond\varphi \vee \Diamond\psi$$

Another important formula which can be seen to be valid is the following

$$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi).$$

This formula (formula scheme to be more precise) is called K (honoring S. Kripke). To see that K is valid, let $\mathcal{M} = (W, R, L)$ be any model and let $x \in W$ be some world in $\mathcal{M}$. Assume that $x \Vdash \Box(\varphi \to \psi)$ and $x \Vdash \Box\varphi$. This holds if and only if for every $y \in W$ such that xRy, we have $y \Vdash \varphi \to \psi$ and $y \Vdash \varphi$ which implies that $y \Vdash \psi$ for every y such that xRy. But this on the other hand holds if and only if $x \Vdash \Box\psi$. Hence $x \Vdash \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, which shows that K is valid.

Exercise 2. Is the converse of the K formula valid? That is, do we have

$$\models (\Box\varphi \to \Box\psi) \to \Box(\varphi \to \psi) \ ?$$

3 Engineering of Modal Logics

As we discussed in the introduction we are interested in modeling situations were we need to distinguish between different modes of truth. And as we saw the applications can vary from temporal to epistemological. The framework for basic modal logic is quite general (although it can be further generalized as we will see later) and can be refined to yield the properties appropriate for the intended application. We will concentrate on three different applications: logic of necessity, temporal logic and logic of knowledge. That is, we will engineer the basic framework to fit the following readings of $\Box\varphi$:

- It is necessarily true that φ
- It will always be true that φ
- Agent A knows φ.

We know that $\Diamond\varphi \equiv \neg\Box\neg\varphi$, so the reading of $\Diamond\varphi$ in each situation is given automatically by that of $\Box\varphi$:

- It is *not* necessarily true that *not* φ
  $\equiv$ It is possible that *not not* φ
  $\equiv$ It is possible that φ.
- It will *not* always be true that *not* φ
  $\equiv$ At some point in the future *not* φ will *not* hold
  $\equiv$ At some point in the future φ will hold.
- Agent A does *not* know *not* φ
  $\equiv$ As far as A knows, φ could be the case
  $\equiv$ φ is consistent with A's knowledge.

Exercise 3. Suppose we want to interpret $\Box\varphi$ as "We have a proof of φ". What would the corresponding interpretation of $\Diamond\varphi$ be?

In the last section we saw some examples of valid formulas, i.e. formulas that are satisfied in every model. Many other formulas, of course, are not. Some examples are $\Box\varphi \to \varphi$, $\Box\varphi \to \Box\Box\varphi$ and $\Diamond\top$ (Why?). However, if we want to study the logic of necessity we would like the first of these, $\Box\varphi \to \varphi$ ('What is necessarily true is also true'), to be valid, in the case that $\Box\varphi$ is read 'Agent A knows φ' we might want $\Box\varphi \to \Box\Box\varphi$ ('If A knows φ, A also knows that he/she/it knows φ') to be valid, and in the case of temporal logic we might want $\Diamond\top$ ('There is always a future world') to be valid.

So for each situation, or reading of $\Box$, we would like to restrict the class of models so that the desired formulas are valid (with respect to this restricted class of models).

Now, each reading of $\Box$ also provides some corresponding reading of the accessibility relation xRy:

- y is a possible world according to the information at x
- y is in the future of x
- y could be the actual world according to A's knowledge at x.

Exercise 4. Consider again the interpretation in exercise 3. Can you say anything about the accessibility relation xRy in this case?

The question now is what properties the relation R should have in the various cases. In the first case for example, is it desirable that R be reflexive? Well, this would mean that at each world x, x itself is a possible world. So the answer seems to be yes. We may note a similarity to the argument validating the formula $\Box\varphi \to \varphi$ under the same reading of $\Box$. In fact, there is a close connection between this formula and the property of reflexivity. In the next section we will see that some elementary classes of models correspond to simple formulas in basic modal logic. This will yield a connection between what formulas should be valid and what general structure the models should have in each situation.

3.1 Frame correspondence

We start with some definitions.

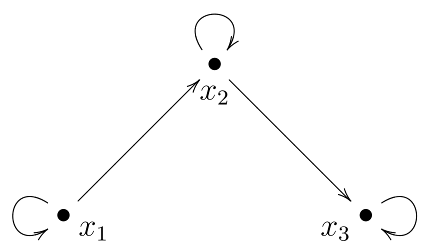
Definition 7. A structure (W, R) with W a non-empty set and R a binary relation on W, is called a *frame* and is denoted $\mathcal{F}$.

A frame $\mathcal{F}$ is the underlying structure of any model $\mathcal{M}$, and so from any model we can extract a frame by simply forgetting about the labeling function.

Definition 8. A formula φ is *valid* on a frame $\mathcal{F}$, written $\mathcal{F} \models \varphi$, if for every labelling function L and each $x \in W$ we have $\mathcal{M}, x \Vdash \varphi$, where $\mathcal{M} = (W, R, L)$.

Remark 3.1. We defined validity of a formula φ, $\models \varphi$, by saying that φ is true at every state of every model, but we could equivalently say that a formula φ is valid when $\mathcal{F} \models \varphi$ for all frames $\mathcal{F}$.

Example 8. Consider the following frame $\mathcal{F}$:



Then we have, $\mathcal{F} \models \Box\varphi \to \varphi$. Why is this? Well, let L be any labelling function on $\mathcal{F}$ and let x be any state of $\mathcal{F}$. If $x \Vdash \Box\varphi$ then $y \Vdash \varphi$ for every y in $\mathcal{F}$ such that x is related to y. But, every state in $\mathcal{F}$ is related to itself, i.e. $\mathcal{F}$ has a reflexive accessibility relation. Hence $x \Vdash \varphi$, and so we indeed have $\mathcal{F} \models \Box\varphi \to \varphi$.

This is a special case of the following result:

Proposition 3.2. *Let $\mathcal{F} = (W, R)$ be a frame, then*

1. *R is reflexive if and only if $\mathcal{F} \models \Box\varphi \to \varphi$,*

2. *R is transitive if and only if $\mathcal{F} \models \Box\varphi \to \Box\Box\varphi$.*

Proof. (1): Suppose R is reflexive and let L be a labelling function on $\mathcal{F}$ so that we get a model $\mathcal{M} = (W, R, L)$. We want to show that $\mathcal{M} \models \Box\varphi \to \varphi$, so let $x \in W$ be any state such that $x \Vdash \Box\varphi$. Since R is reflexive, we have xRx and hence $x \Vdash \varphi$. But then we have $x \Vdash \Box\varphi \to \varphi$, and $\mathcal{F} \models \Box\varphi \to \varphi$ since x was arbitrary.

Conversely, suppose $\mathcal{F} \models \Box\varphi \to \varphi$. In particular, we then have $\mathcal{F} \models \Box p \to p$. Now, let $x \in W$ and let L be a labelling function such that $p \notin L(x)$ and $p \in L(y)$ for each $y \in W$ with xRy. Suppose we don't have xRx, then $x \Vdash \Box p$. But then, since $\mathcal{F}$ satisfies $\Box p \to p$ we also must have $x \Vdash p$. But this is a contradiction to the assumption that $p \notin L(x)$. Hence, it must be the case that xRx. Since x was arbitrary this shows that R is reflexive.

(2): Suppose R is transitive. Let L be a labelling function and $\mathcal{M} = (W, R, L)$. We want to show that $\mathcal{M} \models \Box\varphi \to \Box\Box\varphi$. So let $x \in W$ be any state such that $x \Vdash \Box\varphi$. We then need to show that for every $y \in W$ such that xRy and every $z \in W$ such that yRz we have $z \Vdash \varphi$, i.e. $x \Vdash \Box\Box\varphi$. But if xRy and yRz then xRz since R is transitive, and together with $x \Vdash \Box\varphi$ we then have $z \Vdash \varphi$. Hence $x \Vdash \Box\Box\varphi$. This shows that $\mathcal{F} \models \Box\varphi \to \Box\Box\varphi$.

Conversely, suppose $\mathcal{F} \models \Box\varphi \to \Box\Box\varphi$, In particular, we then have $\mathcal{F} \models \Box p \to \Box\Box p$. Let $x, y, z \in W$ be such that xRy and yRz, we want to show xRz. Let L be a labelling function such that $p \notin L(z)$ but $p \in L(w)$ for all other worlds w. Suppose we don't have xRz, then $x \Vdash \Box p$ and hence $x \Vdash \Box\Box p$ since $\mathcal{F} \models \Box p \to \Box\Box p$. But then $y \Vdash \Box p$, since xRy, and $z \Vdash p$, since yRz, which contradicts our assumption that $p \notin L(z)$. Hence, we must have xRz. This shows that R is transitive. $\Box$

For other applications there might be other properties of R that are desirable. And in many cases these properties will, as above, correspond to some formula. The following table gives some such correspondences

T: Frame-validity of $\Box\varphi \to \varphi$ corresponds to reflexivity of R.

B: Frame-validity of $\varphi \to \Box\Diamond\varphi$ corresponds to symmetry of R.

D: Frame-validity of $\Box\varphi \to \Diamond\varphi$ corresponds to R being serial.

4: Frame-validity of $\Box\varphi \to \Box\Box\varphi$ corresponds to transitivity of R.

5: Frame-validity of $\Diamond\varphi \to \Box\Diamond\varphi$ corresponds to R being Euclidean.

The first symbol on each line is the commonly used name of the corresponding modal formula (cf. axiom K).

Definition 9. We say that a BML formula φ *defines* a property P of a frame $\mathcal{F} = (W, R)$ if $\mathcal{F} \models \varphi$ if and only if R has the property P.

Exercise 5. Show that both $\Diamond\top$ and $D : \Box\varphi \to \Diamond\varphi$ defines the same property.

Example 9. We will prove the last of the correspondences above, that $\mathcal{F} \models \Diamond\varphi \to \Box\Diamond\varphi$ if and only if R is Euclidean. The relation R is Euclidean if for every $x, y, z \in W$, xRy and xRz implies that yRz. First, suppose that R is Euclidean. Let L be any labelling function on $\mathcal{F}$ and let $x \in W$ such that $x \Vdash \Diamond\varphi$. Then there is $z \in W$ with xRz and $z \Vdash \varphi$. Now suppose $y \in W$ with xRy, then yRz since R is Euclidean. But then we have $y \Vdash \Diamond\varphi$, and hence $x \Vdash \Box\Diamond\varphi$, i.e. $x \Vdash \Diamond\varphi \to \Box\Diamond\varphi$.

We prove the converse by contraposition. Assume $\mathcal{F}$ is non-Euclidean, then there must be states $x, y, z \in W$ such that xRy, xRz but not yRz. We will try to falsify 5 in x by finding a labelling function L such that $x \Vdash \Diamond p$ and $x \nVdash \Box\Diamond p$. That is, we have to make p true at some R-successor of x and false at all R-successors of some R-successor of x. Let L be given by

$$p \in L(w) \text{ iff it is not the case that } yRw$$

then $p \in L(z)$ while $\{w \mid yRw\} \cap \{w \mid p \in L(w)\} = \emptyset$. Now clearly $y \nVdash \Diamond p$, so that $x \nVdash \Box\Diamond p$. On the other hand, since we have $z \Vdash p$ and xRz, we have $x \Vdash \Diamond p$. Hence, $\mathcal{F} \not\models \Diamond\varphi \to \Box\Diamond\varphi$.

Exercise 6. Prove the remaining frame correspondences in the list.

Exercise 7. Can you find a modal formula that defines linearity? R is *linear* if it is reflexive, transitive and satisfies $(\forall x, y)(xRy \vee yRx)$.

We now have a way of deciding what formulas of basic modal logic should be included as axioms in our logic: On the one hand we are guided by the reading of the unary connectives $\Box$ and $\Diamond$, and on the other hand by the desired properties of models.

For example, say we want to interpret $\Box$ as the temporal connective *Always in the future*. Then we have already argued that we would like to have the formula $\Diamond\top$ as an axiom. Furthermore it would be reasonable to consider only transitive models, which would simply mean that if y is ahead in time of x and z is ahead in time of y, then z is also ahead in time of x. So we add the formula 4 as an axiom.

So how could logics for our three readings of $\Box\varphi$ look like? Before we can investigate this further we need a proper definition of what we mean by a *logic*.

3.2 Normal modal logics

Given a class of frames $\mathbf{F}$, we denote by $\Lambda_{\mathbf{F}}$ the set of formulas valid on every frame in $\mathbf{F}$. So for example, if $\mathbf{F}$ is the class of all reflexive frames, we know that $\Box\varphi \rightarrow \varphi \in \Lambda_{\mathbf{F}}$. Now, are there syntactic mechanisms capable of generating $\Lambda_{\mathbf{F}}$? And are such mechanisms able to cope with the associated semantic consequence relation?

We are going to define a Hilbert-style axiom system called **K**, which is a 'minimal' system for reasoning about frames.

Definition 10. A **K**-*proof* is a finite sequence of formulas, each of which is an *axiom*, or follows from one or more earlier items in the sequence by applying *a rule of proof*. The axioms of **K** are all instances of propositional tautologies and:

K: $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$,

Dual: $\Diamond\varphi \leftrightarrow \neg\Box\neg\varphi$.

The rules of proof of **K** are:

- *Modus ponens*: given φ and $\varphi \rightarrow \psi$, prove ψ.
- *Uniform substitution*: given φ, prove ψ, where ψ is obtained from φ by replacing proposition letters in φ by arbitrary formulas.

- *Rule of necessitation*: given φ, prove $\Box\varphi$.

A formula φ is **K**-*provable* if it occurs as the last item of some **K**-proof, in this case we write $\vdash_{\mathbf{K}} \varphi$.

The definition needs some explaining. Adding *all* propositional tautologies of course yields a very large axiom set and we could have chosen a small set of tautologies capable of generating the rest by using the rules of proof. However, we are not at the moment interested in having a minimal generating set of axioms.

Modus ponens preserves validity. That is, if $\models \varphi$ and $\models \varphi \to \psi$ then also $\models \psi$ so it is a correct rule for reasoning about frames. Furthermore, modus ponens preserves *global truth* (if $\mathcal{M} \models \varphi$ and $\mathcal{M} \models \varphi \to \psi$ then also $\mathcal{M} \models \psi$) and *satisfiability* (if $\mathcal{M}, x \Vdash \varphi$ and $\mathcal{M}, x \Vdash \varphi \to \psi$ then also $\mathcal{M}, x \Vdash \psi$). Thus, modus ponens is also a correct rule for reasoning about models, both globally and locally.

Uniform substitution mirrors the fact that validity has nothing to do with particular assignments: if a formula is valid, this does not depend on the particular values its propositional symbols have, thus we should be free to uniformly replace any propositional symbol with an arbitrary formula. Uniform substitution preserves validity (why?), but neither global truth nor satisfiability. For example q is obtainable from p by uniform substitution but just because p is globally true in some model it does not follow that the same is true for q.

The K axiom lets us transform a boxed formula $\Box(\varphi \to \psi)$ into an implication $\Box\varphi \to \Box\psi$. It is sometimes called the *distribution axiom*. And as we have already seen, K is valid on all frames.

The reason for having the Dual axiom is that we did not define $\Diamond$ using box. We saw earlier that it is a valid formula scheme.

The rule of necessitation might look somewhat odd, since clearly $\varphi \to \Box\varphi$ is not valid. However, the rule of necessitation preserves validity; if φ is valid, then also $\Box\varphi$ is valid. Similarly, it preserves *global truth*; if $\mathcal{M} \models \varphi$ then $\mathcal{M} \models \Box\varphi$ (why?).

Exercise 8. Show that all the rules above preserve validity. Which of the rules preserves global and/or local truth?

K is the minimal modal Hilbert system in the following sense: All its axioms are valid and all the rules of inference preserve validity, hence all **K**-provable formulas are valid. That is **K** is sound with respect to the class of all frames. Moreover, the converse is also true: if a formula of basic modal logic is valid, then it is **K**-provable. The proof of this fact is way beyond the scope of the present presentation.

Example 10. The formula $(\Box p \wedge \Box q) \to \Box(p \wedge q)$ is valid on any frame, so it should be **K**-provable, and indeed it is. Consider the following sequence of formulas

Tautology

1. $\vdash p \to (q \to (p \wedge q))$

Generalization: 1

2. $\vdash \Box(p \to (q \to (p \wedge q)))$

K axiom

3. $\vdash \Box(p \to q) \to (\Box p \to \Box q)$

Uniform Substitution: 3

4. $\vdash \Box(p \to (q \to (p \wedge q))) \to (\Box p \to \Box(q \to (p \wedge q)))$

Modus Ponens: 2,4

5. $\vdash \Box p \to \Box(q \to (p \wedge q))$

Uniform Substitution: 3

6. $\vdash \Box(q \to (p \wedge q)) \to (\Box q \to \Box(p \wedge q))$

Propositional Logic: 5,6

7. $\vdash \Box p \to (\Box q \to \Box(p \wedge q))$

Propositional Logic: 7

8. $\vdash (\Box p \wedge \Box q) \to \Box(p \wedge q)$

As a matter of fact we have cheated a bit here; some of the steps are missing. So this is not a **K**-proof in the strict sense, although we see that it is possible to fill in the gaps (from 6 to 8) in order to get a complete proof.

Suppose now that we are interested in validity only on transitive frames. Then we know the formula $\Box\varphi \to \Box\Box\varphi$ is valid on this class of frames, and hence we would like to be able to derive it. But the system **K** is to weak for this, since it only derives valid formulas (that is, formulas valid on all frames) and $\Box\varphi \to \Box\Box\varphi$ is not valid. However, we can simply add $\Box\varphi \to \Box\Box\varphi$ to **K** as an axiom, we then obtain the Hilbert system **K4**. It is then possible to show that **K4** is sound and complete with respect to the class of all transitive frames. That is, **K4** generates precisely the formulas valid on transitive frames. More generally we may add any set of modal formulas Γ as axioms to **K** and obtain an axiom system **K**Γ.

We will now introduce the concept of a normal modal logic.

Definition 11. A *normal modal logic* L is a set of formulas of basic modal logic, with the following properties:

(1) L contains all tautologies

(2) L contains all instances of the formula scheme K:

$$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$$

(3) L contains all instances of the formula scheme Dual:

$$\Diamond\varphi \leftrightarrow \neg\Box\neg\varphi$$

(4) L is closed under uniform substitution and modus ponens.

(5) L is closed under the rule of necessitation.

We call the smallest normal modal logic **K**, and it just contains propositional logic and all instances of axiom K, together with all formulas that you get by applying conditions (3)-(5) above.

Now, to "build" a modal logic, first choose the formula schemes that you would like to have in it, these will be the *axioms* of the logic. Then close it under the conditions of the definition.

In the case $\Box\varphi$ is read 'It is necessarily true that φ' we may (as we discussed earlier) want $\Box\varphi \to \Diamond\varphi$ and $\varphi \to \Diamond\varphi$ to be in our logic L. Then we can see, by frame correspondence, that every model of L will have a serial accessibility relation R. Moreover, R must be reflexive since $\varphi \to \Diamond\varphi \equiv \Box\varphi \to \varphi$. A possible name for L could then be **KTD**.

If we instead look at the case when $\Box\varphi$ is read 'It will always be true that φ'. In this case we may, as we discussed earlier, want to add the formula $\Diamond\top$. Moreover we may want the present to be part of the future and therefore add $\Box\varphi \to \varphi$. Then we get a logic, whose models are reflexive and satisfy for every $x \in W$ there is $y \in W$ such that xRy. We may of course want further refinements depending on the situation we want to model.

The last case we will look at in some more detail.

Example 11 (The modal Logic **KT45**)**.** Here $\Box\varphi$ is read 'Agent A knows φ'. A logic commonly used in this situation is **KT45**, which means that we add to **K** the formula schemes $T : \Box\varphi \to \varphi$, $4 : \Box\varphi \to \Box\Box\varphi$ and $5 : \Diamond\varphi \to \Box\Diamond\varphi$, and close under the conditions of the definition. The axioms T, 4 and 5 tell us that

T. Truth: the agent A knows only true things.

4. Positive introspection: if the agent A knows something, then he/she/it knows that he/she/it knows it.

5. Negative introspection: if the agent A doesn't know something, then he/she/it knows that he/she/it doesn't know it.

The K axiom tells us that the agents knowledge is closed under logical consequence. These properties represent idealizations of knowledge, and **KT45** might not be appropriate in some, or even many, situations.

What we do know, is that the semantics for **KT45** must consider only frames where the accessibility relation is reflexive (T), transitive (4) and Euclidean (5). In fact one can prove that such a relation must be an equivalence relation, i.e. reflexive, transitive and symmetric. Hence we know that $B : \varphi \to \Box\Diamond\varphi$ is true in every model of **KT45**.

Exercise 9. Show that in a frame, $\mathcal{F} = (W, R)$, for **KT45** the accessibility relation R must be an equivalence relation.

4 Neighborhood Semantics: A remark on normal modal logics

When we introduced the Hilbert system **K**, the motivation for having the formula

$$\text{(K):} \qquad \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi),$$

as an axiom was simply its validity with respect to the class of all frames.

Removing (K) from the set of axioms would also yield a type of modal logic. However, then the Kripke style semantics would no longer be the right semantics (simply because of the validity of (K) in this semantics).

One possible semantics, in which (K) is no longer valid, is called *neighborhood semantics*.

Definition 12. A *neighborhood frame* is a pair (W, N), where W is a set and N is a map

$$N : W \to \mathcal{P}(\mathcal{P}(W)).$$

N is called a *neighborhood function* and assigns to each $w \in W$ a set $N(w)$ of neighborhoods of w.

A model is as before a frame together with a labelling of the states with atoms.

Definition 13. A *neighborhood model* is a neighborhood frame (W, N) together with a *labelling function*

$$L : W \to \mathcal{P}(Atoms).$$

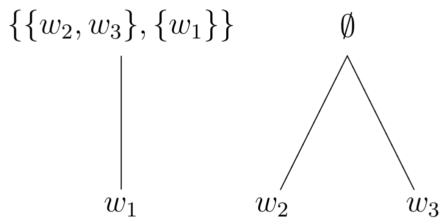
Now, the truth of boxed formulas $\Box\varphi$ at a state w will be interpreted as φ being true in a neighborhood of w.

Definition 14. We define that a formula φ of BML is true at a state w in a model $\mathcal{M} = (W, N, L)$ by induction on φ, and denote this as $\mathcal{M}, w \Vdash \varphi$ (or rather $w \Vdash \varphi$ when the model $\mathcal{M}$ is understood),

- $x \Vdash \top$
- $x \nVdash \bot$
- $x \Vdash p$ iff $p \in L(x)$
- $x \Vdash \neg\varphi$ iff $x \nVdash \varphi$
- $x \Vdash \varphi \wedge \psi$ iff $x \Vdash \varphi$ and $x \Vdash \psi$
- $x \Vdash \varphi \vee \psi$ iff $x \Vdash \varphi$ or $x \Vdash \psi$
- $x \Vdash \varphi \to \psi$ iff $x \Vdash \psi$ whenever $x \Vdash \varphi$
- $x \Vdash \varphi \leftrightarrow \psi$ iff $x \Vdash \varphi \Leftrightarrow x \Vdash \psi$
- $x \Vdash \Box\varphi$ iff $\varphi^{\mathcal{M}} \in N(x)$
- $x \Vdash \Diamond\varphi$ iff $W \setminus \varphi^{\mathcal{M}} \notin N(x)$

Here $\varphi^{\mathcal{M}} := \{y \in W \mid \mathcal{M}, y \Vdash \varphi\}$ and is called the *truth set* of φ.

Example 12. Consider the neighborhood model $\mathcal{M} = (W, N, L)$ with $W = \{w_1, w_2, w_3\}$, $N(w_1) = \{\{w_2, w_3\}, \{w_1\}\}$, $N(w_i) = \emptyset$ for $i = 2, 3$ and $L(w_1) = \{p\}$, $L(w_2) = \{q\}$, $L(w_3) = \emptyset$.



Then $(p \to q)^{\mathcal{M}} = \{w \mid w \Vdash p \to q\} = \{w_2, w_3\} \in N(w_1)$ and hence we have $w_1 \Vdash \Box(p \to q)$. We also have $w_1 \Vdash \Box p$, since $p^{\mathcal{M}} = \{w_1\} \in N(w_1)$. But $w_1 \nVdash \Box q$ since $q^{\mathcal{M}} = \{w_2\} \notin N(w_1)$, and we have

$$w_1 \nVdash \Box(p \to q) \to (\Box p \to \Box q).$$

So (K) is not satisfied in $\mathcal{M}$.

Exercise 10. Show that to every Kripke model $\mathcal{K} = (M, R, L)$ there corresponds an equivalent neighborhood model $\mathcal{N} = (W, N, L)$ where $N : W \to \mathcal{P}(\mathcal{P}(W))$ given by

$$N(w) := \{\varphi^{\mathcal{K}} \mid \mathcal{K}, w \Vdash \Box\varphi\}.$$

(By equivalent we mean that for all $w \in W$: $\mathcal{K}, w \Vdash \psi$ iff $\mathcal{N}, w \Vdash \psi$).

Is the converse true? That is, given a neighborhood model $\mathcal{N}$ can we find an equivalent Kripke model $\mathcal{K}$?

Thus, neighborhood semantics generalizes the possible worlds semantics.

As before we may speak of frame correspondence.

Definition 15. We say that a BML formula φ *defines* a property P of a neighborhood frame $\mathcal{F} = (W, N)$ if $\mathcal{F} \models \varphi$ if and only if N has the property P.

Recall that on Kripke frames the two formulas $\Diamond\top$ and $\Box\varphi \to \Diamond\varphi$ define the same property: seriality. For neighborhood frames this is no longer the case.

Lemma 1. *Let $\mathcal{F} = (W, N)$ be a neighborhood frame. Then*

(i) $\mathcal{F} \models \Diamond\top$ *if and only if* $\emptyset \notin N(w)$ *for all* $w \in W$.

(ii) $\mathcal{F} \models \Box\varphi \to \Diamond\varphi$ *if and only if* $N(w)$ *is* proper *(i.e. If* $U \in N(w)$ *implies* $U^c \notin N(w)$*).*

Proof. (i) follows directly from the definition of truth for formulas $\Diamond\varphi$.

(ii) Suppose $N(w)$ is proper for every $w \in W$ and let L be an arbitrary labelling function on $\mathcal{F}$. Let $w \in W$ such that $w \Vdash \Box\varphi$, then $\varphi^{\mathcal{M}} \in N(w)$. Hence $W \setminus \varphi^{\mathcal{M}} = (\varphi^{\mathcal{M}})^c \notin N(w)$ and we have $w \Vdash \Diamond\varphi$. Hence $\mathcal{F} \models \Box\varphi \to \Diamond\varphi$.

Conversely, suppose $\mathcal{F} \models \Box\varphi \to \Diamond\varphi$ and that for some $w \in W$ there is $U \in N(w)$ such that also $U^c \in N(w)$. Define a labelling function L by setting $p \in L(x)$ if and only if $x \in U$. Then $p^{\mathcal{M}} = U$ and we have $w \Vdash \Box p$. But since $W \setminus p^{\mathcal{M}} = U^c \in N(w)$ we have $w \nVdash \Diamond p$. This contradicts the assumption that $\mathcal{F} \models \Box\varphi \to \Diamond\varphi$. We conclude that $N(w)$ is proper for all $w \in W$. $\Box$

Exercise 11. Find properties that are defined by the following formulas:

(1) $\Box\varphi \to \varphi$,

(2) $\Box\varphi \to \Box\Box\varphi$,

(3) $\Box\bot$,

(4) $\neg\Box\varphi \to \Box\neg\Box\varphi$,

(5) $\Diamond\varphi \to \Box\varphi$.

5 Intuitionistic Propositional Calculus

We will in this section see how modal logic provides a semantics for intuitionistic propositional logic (IPC).

Recall that IPC is a propositional logic without the rule of reductio ad absurdum, i.e. $\neg\neg\varphi \to \varphi$ is not derivable. We will actually prove this in a moment, but first we need to define the syntax and a proof system for IPC.

Definition 16. Formulas φ of IPC are given by the following rule

$$\varphi ::= \bot \mid p \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \to \varphi$$

Where p is an arbitrary propositional symbol in *Atoms* - the set of propositional symbols. Negation $\neg\varphi$ is defined as $\varphi \to \bot$.

A system of natural deduction IPC is given by:

(A) $\quad \varphi_1, \ldots, \varphi_n \vdash \varphi_i$

$(\bot)$ $$\frac{\Gamma \vdash \bot}{\Gamma \vdash \varphi}$$

$(\wedge I)$ $$\frac{\Gamma \vdash \varphi \quad \Gamma \vdash \psi}{\Gamma \vdash \varphi \wedge \psi}$$

$(\wedge E)$ $$\frac{\Gamma \vdash \varphi \wedge \psi}{\Gamma \vdash \varphi} \qquad \frac{\Gamma \vdash \varphi \wedge \psi}{\Gamma \vdash \psi}$$

$(\vee I)$ $$\frac{\Gamma \vdash \varphi}{\Gamma \vdash \varphi \vee \psi} \qquad \frac{\Gamma \vdash \psi}{\Gamma \vdash \varphi \vee \psi}$$

$(\vee E)$ $$\frac{\Gamma \vdash \varphi \vee \psi \quad \Gamma, \varphi \vdash \chi \quad \Gamma, \psi \vdash \chi}{\Gamma \vdash \chi}$$

$(\to I)$ $$\frac{\Gamma, \varphi \vdash \psi}{\Gamma \vdash \varphi \to \psi}$$

$(\to E)$ $$\frac{\Gamma \vdash \varphi \quad \Gamma \vdash \varphi \to \psi}{\Gamma \vdash \psi}$$

Where $\Gamma \vdash \varphi$ means that φ is derivable under assumptions $\Gamma = \varphi_1, \ldots, \varphi_n$. When Γ is empty we simply write $\vdash \varphi$.

Example 13. We show $\vdash \varphi \to (\neg\varphi \to \psi)$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\varphi, \neg\varphi \vdash \varphi \quad \varphi, \neg\varphi \vdash \neg\varphi}{\varphi, \neg\varphi \vdash \bot} \to E}{\varphi, \neg\varphi \vdash \psi} \bot}{\varphi \vdash \neg\varphi \to \psi} \to I}{\vdash \varphi \to (\neg\varphi \to \psi)} \to I$$

Exercise 12. Give a derivation for $\vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$.

Even though IPC does not have any modal operators, we can give meaning to the logical constants with a Kripke style semantics. As in the case of modal logics, a model for IPC will consist of a frame (W, R) and a labelling function $L : W \to \mathcal{P}(Atoms)$. We will think of the elements of W as 'information states' or 'bundles of data', that can be incomplete in the sense that the collected data at some state is possibly not enough to decide the truth value of every statement that can be expressed in the language. Compare this to the classical situation where every statement is either true or false at any state.

The accessibility relation R is going to be a reflexive partial order and we will read iRj as

"Information state j can still be reached once the information of state i is already acquired"

We want to think of states iRj as j being a state where we have acquired more (or at least the same) knowledge than in i. For this to hold formally we need a forcing relation $\Vdash$ satisfying

$$iRj \Longrightarrow (i \Vdash \varphi \Longrightarrow j \Vdash \varphi).$$

We can only enforce this for atoms in our definition, but we will see that the property carries over to arbitrary formulas of IPC.

Definition 17. A model for IPC is a Kripke model (W, R, L) such that

- (i) R is a reflexive partial order and,
- (ii) L satisfies: If iRj then $L(i) \subseteq L(j)$.

Truth at a state in a model for IPC is defined as usual by an induction on formulas:

- $i \nVdash \bot$,

- $i \Vdash \varphi \wedge \psi$ iff $i \Vdash \varphi$ and $i \Vdash \psi$,
- $i \Vdash \varphi \vee \psi$ iff $i \Vdash \varphi$ or $i \Vdash \psi$,
- $i \Vdash \varphi \to \psi$ iff for all $j \in W$ such that iRj we have $j \nVdash \varphi$ or $j \Vdash \psi$.

Exercise 13. Show that $i \Vdash \neg\varphi$ if and only if $j \nVdash \varphi$ for all j such that iRj. Remember that $\neg\varphi \equiv \varphi \to \bot$.

We note that $i \nVdash \varphi$ only means that no verification of φ have been found *yet* as opposed to $i \Vdash \neg\varphi$ meaning that we will never be able to find one.

We will now prove that knowledge or truth is preserved by the relation R.

Lemma 2. *Let $i, j \in W$ and φ be a formula of IPC. Then $i \Vdash \varphi$ and iRj implies $j \Vdash \varphi$.*

Proof. By induction on the complexity of φ.

The base case, i.e. $\varphi = p$ atomic, follows by the definition of a model of IPC. The cases $\varphi = \bot$ and $\varphi = \varphi_1 \wedge \varphi_2$ are straightforward and so we jump directly to the interesting case: $\varphi = \varphi_1 \to \varphi_2$.

Suppose $i \Vdash \varphi_1 \to \varphi_2$, then $k \nVdash \varphi_1$ or $k \Vdash \varphi_2$ for all k such that iRk. But then $k \nVdash \varphi_1$ or $k \Vdash \varphi_2$ for all k such that jRk since iRj and R is transitive. But this just means that $j \Vdash \varphi_1 \to \varphi_2$. □

Exercise 14. Fill in the gaps in the above proof.

As usual we would like our deduction system to be *sound*, i.e. it should only be able to derive true formulas (with respect to the semantics just given). If $\Gamma = \varphi_1, \ldots, \varphi_n$, then we mean by $\mathcal{M} \models \Gamma \to \varphi$, that $\mathcal{M} \models (\varphi_1 \wedge \cdots \wedge \varphi_n) \to \varphi$.

Theorem 5.1 (Soundness)**.** *The semantics given above is sound. That is, if $\mathcal{M} = (W, R, L)$ is a model for IPC, then*

$$\Gamma \vdash \varphi \Longrightarrow \mathcal{M} \models \Gamma \to \varphi.$$

Proof. We use induction on the length of the derivation of $\Gamma \vdash \varphi$.

The base case $\varphi \in \Gamma$ is trivial, and the cases involving $\wedge$ and $\vee$ are straightforward so we will only address the case of implication introduction $(\to I)$

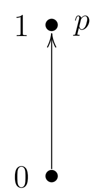
and leave the rest as an exercise. Suppose therefore we have $\mathcal{M} \models \Gamma, \varphi \to \psi$. Let i in $\mathcal{M}$ be arbitrary such that $i \Vdash \Gamma$, i.e. $i \Vdash \varphi_1 \wedge \cdots \wedge \varphi_n$. Then let j be such that iRj and $j \Vdash \varphi$. By lemma (2) we have that also $j \Vdash \Gamma$ and so $j \Vdash \psi$. But this just means that $i \Vdash \varphi \to \psi$, i.e. $\mathcal{M} \models \Gamma \to (\varphi \to \psi)$. □

Exercise 15. Fill in the gaps in the above proof.

The soundness theorem can be used to show that certain formulas of IPC are not derivable by the construction of models not satisfying them. This is just the contrapositive statement of the theorem:

$$\mathcal{M} \not\models \Gamma \to \varphi \Longrightarrow \Gamma \nvdash \varphi.$$

Example 14. $\vdash p \vee \neg p$ is not derivable in IPC. Consider the Kripke model $\mathcal{M}$:



That is $0R1$, $1 \Vdash p$ and $0 \nVdash p$. Then $0 \Vdash p \vee \neg p$ if and only if $0 \Vdash p$ or $o \Vdash \neg p$. But $0 \nVdash p$ and $0 \Vdash \neg p$ would imply $1 \nVdash p$. Hence we must have $0 \nVdash p \vee \neg p$. Therefore $\mathcal{M} \not\models p \vee \neg p$.

Similarly we can show that $\nvdash \neg\neg p \to p$. Consider again the model $\mathcal{M}$ above. We have $0 \Vdash \neg\neg p$, i.e. $0 \Vdash \neg p \to \bot$, since $0 \nVdash \neg p$ and $1 \nVdash \neg p$. But $0 \nVdash p$, so $0 \nVdash \neg\neg p \to p$. Therefore $\mathcal{M} \not\models \neg\neg p \to p$.

Exercise 16. Show that $\nvdash (p \to q) \vee (q \to p)$ in IPC.

Remark 5.2. One can prove that IPC is complete with respects to the Kripke semantics given above, and moreover that it is complete with respect to the class of all finite Kripke models, i.e. IPC has the *finite model property*: If $\nvdash \varphi$ then there is a finite model $\mathcal{M}$ such that $\mathcal{M} \not\models \varphi$.

We will now describe the connection to the framework of basic modal logic. This is done via a translation $(\cdot)^* : IPC \to BML$, defined by

- $p^* = \Box p$, for all $p \in Atoms$,

- $\bot^* = \bot$
- $(\varphi \wedge \psi)^* = \varphi^* \wedge \psi^*$,
- $(\varphi \vee \psi)^* = \varphi^* \vee \psi^*$,
- $(\varphi \to \psi)^* = \Box(\varphi^* \to \psi^*)$.

Note here that in the last bullet the implication used on the left hand side is the implication of IPC while the implication on the right hand side is the one of BML.

The frames, (W, R), for IPC has a reflexive and transitive accessibility relation R, so we could guess that the right semantics for IPC* is something like the one for the normal modal logic **KT4**, i.e. the normal modal logic **K** with the formula (T), corresponding to reflexivity, and the formula (4), corresponding to transitivity, added as axioms. That is, we guess that the class of models appropriate for **KT4** is also appropriate for IPC*. As a matter of fact we have.

Proposition 5.3. $\vdash_{IPC} \varphi$ *if and only if* $\vdash_{\mathbf{KT4}} \varphi^*$.

Proof. We will only sketch a proof of the right to left direction. The other direction is left, since a proof would take us beyond the scope of these notes.

$(\Leftarrow)$: Suppose $\nvdash_{IPC} \varphi$, then there is a finite model (W, R, L) such that for some $i_0 \in W$, $i_0 \nVdash \varphi$. Now, (W, R, L) can be considered a model of BML, where we have a forcing relation $\Vdash^*$ for the BML language (i.e. $i \Vdash^* \psi$, for BML formulas ψ, is defined as in section 2.2).

We claim that $i \Vdash^* \psi^*$ if and only if $i \Vdash \psi$ for all IPC formulas ψ. Then $i_0 \nVdash \varphi$ implies $i_0 \nVdash^* \varphi^*$, and we have $\nvdash_{\mathbf{KT4}} \varphi^*$.

The claim is proved using induction on the complexity of ψ, and as always we only prove one case and leave the rest as exercise. Suppose $\psi = \psi_1 \to \psi_2$ and $i \Vdash \psi$, then for all j such that iRj we have that $j \Vdash \psi_1$ implies $\jmath \Vdash \psi_2$. But this is just $j \Vdash^* \psi_1^* \to \psi_2^*$ for all j with iRj, that is $i \Vdash \Box(\psi_1^* \to \psi^*)$, where $\Box(\psi_1^* \to \psi^*) = \psi^*$. The converse direction is similar. $\Box$

Now, since **KT4** is complete with respect to the class of all reflexive and transitive frames (chapter 4.3 in [1]) the same is true for IPC*.

6 Generalizing the basic framework

So far our three examples have been quite simple. For example, in the case of temporal logic we only had the possibility to talk about truth in the future while it is natural to also want to be able to talk about truth in the past. In the case where we read $\Box\varphi$ as 'agent A knows φ', we could only handle one agent, whereas in a practical situation we would like to be able to model a situation with more than one agent. So let us generalize the basic framework we have developed so far. First, there seems to be no good reason to restrict ourselves to languages with only one 'box'. Second, there seems no good reason to restrict ourselves to unary modalities.

Definition 18. A *modal similarity type* is a pair $\tau = (O, \rho)$ where O is a non-empty set, and ρ is a function $O \to \mathbb{N}$. The elements of O are called *modal operators*; and are denoted $\nabla_0, \nabla_1, \ldots$. The function ρ assigns to each operator $\nabla \in O$ a finite arity, indicating the number of arguments ∇ can be applied to. We keep the word box for unary operators and denote them $\Box_i$ or $[\ i\ ]$ for i in some index set.

Definition 19. A *modal language* is now given by a modal similarity type $\tau = (O, \rho)$ and a set *Atoms* of propositional letters. A formula in this language is given by the rule:

$$\varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \to \varphi \mid \nabla(\varphi, \ldots, \varphi)$$

where the number of arguments that ∇ takes is $\rho(\nabla)$ and $p \in \mathit{Atoms}$.

The dual $\triangle$ of ∇ is defined as $\triangle(\varphi_1, \ldots, \varphi_n) := \neg\nabla(\neg\varphi_1 \ldots, \neg\varphi_n)$, when $\rho(\nabla) = n$. The dual of a box is called a diamond, and is denoted $\Diamond_i$ or $\langle i \rangle$

Models will now have to encompass an accessibility relation for each modal operator.

Definition 20. Let τ be a modal similarity type. A τ-frame is a tuple $\mathcal{F}$ consisting of

(i) A set W of worlds

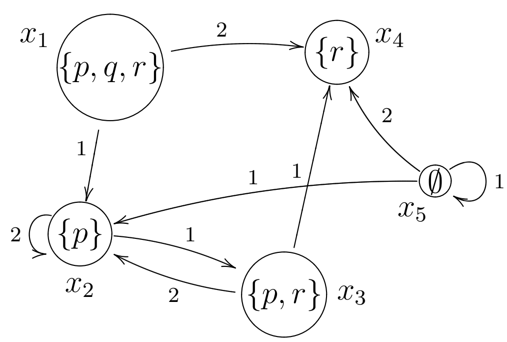
(ii) for each $n \geq 0$, and each n-ary modal operator ∇ in the similarity type τ an $(n+1)$-ary relation R_∇.

A τ-model $\mathcal{M}$ is simply a τ-frame $\mathcal{F}$ together with a labelling function L, that is $\mathcal{M} = (\mathcal{F}, L)$.

The notion of a formula φ being satisfied in a world x in a model $\mathcal{M} = (W, \{R_\nabla \mid \nabla \in \tau\}, L)$, denoted $\mathcal{M}, x \Vdash \varphi$ is defined inductively. The only case different from basic modal logic being the modal case:

$$\mathcal{M}, x \Vdash \nabla(\varphi_1, \ldots, \varphi_n) \quad \text{iff} \quad \begin{array}{l}\text{for every } y_1, \ldots, y_n \in W \text{ with } (x, y_1, \ldots, y_2) \in R_\nabla \\ \text{we have for each } i \text{ that } \mathcal{M}, y_i \Vdash \varphi_i.\end{array}$$

Example 15. If we have a modal language with two unary modalities $\Box_1$ and $\Box_2$, a model will have two corresponding binary relations R_1 and R_2 and could look as follows:



So the underlying frame is a *labeled transition system*

Example 16 (The Basic Temporal Language)**.** The basic temporal language (BTL) is built using a set of unary operators $O = \{[G], [H]\}$ where the intended interpretation of a formula $[G]\varphi$ is 'φ will always be true in the future' and the intended interpretation of a formula $[H]\varphi$ is 'φ has always been true in the past'. Their dual are denoted $\langle F \rangle$ and $\langle P \rangle$ respectively. With this language we can express many more things than with our previous temporal language. For example: $\langle P \rangle \varphi \to [G]\langle P \rangle \varphi$ says that 'whatever has happened will always have happened'. We will often write just the boxed letter to denote the modality, e.g. G instead of $[G]$.

Since we now have two unary operators, a model for the language will have to be based on a frame with two binary relations, R_G and R_H. But since we are interested in modeling time we don't want them to be any two binary relations, rather we want them to be converse. That is, xR_Gy if and only if yR_Hx. Let us denote the converse of a relation R by $\check{R}$. We will call a model of the form $(W, R, \check{R}, L)$ a *bidirectional model*, and similarly the underlying frame is called a *bidirectional frame*. In this case we usually write $\mathcal{F} = (W, R)$

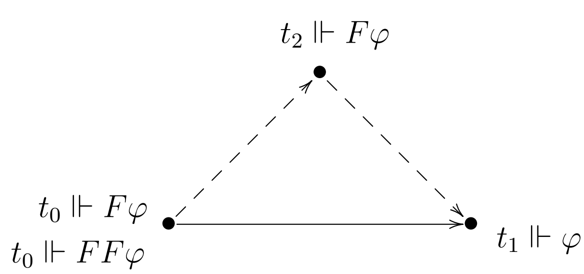
Exercise 17. Try to give a formula characterizing bidirectional frames.

Example 17. Consider again the the basic temporal language, and suppose we want to consider *dense* (*bidirectional*) frames $\mathcal{F} = (W, R)$. That is, R satisfies

$$(\forall x, y \in W)(xRy \to (\exists z \in W)(xRz \wedge zRy)).$$

Can we characterize this property by a formula in BTL?

We may think as follows: If the formula $F\varphi$ holds at time t_0, then there is a time t_1 such that t_0Rt_1 and where φ holds. If the frame is dense, we should be able to find t_2 such that t_0Rt_2 and t_2Rt_1, and hence also $FF\varphi$ should hold at t_0.



So let us try with the formula

$$F\varphi \to FF\varphi.$$

Suppose $\mathcal{F}$ is dense and suppose $t \Vdash Fp$ under some arbitrary labelling function L. Then there is a state $t' \in W$ such that tRt' and $t' \Vdash p$. But as $\mathcal{F}$ is dense there is $s \in W$ such that tRs and sRt'. So $s \Vdash Fp$ and $t \Vdash FFp$.

Conversely, suppose that $\mathcal{F}$ is a frame such that $\mathcal{F} \models F\varphi \to FF\varphi$, and suppose $t \in W$ has an R-successor t'. Let L be the minimal labelling function defined by

$$L(t') = \{p\}.$$

Then we have $\mathcal{M}, t \Vdash Fp$, where $\mathcal{M} = \mathcal{F}, L = (W, R, L)$. Since $\mathcal{F} \models Fp \to FFp$ we must have $t \Vdash FFp$. This means there is a state $s \in W$ such that tRs and $s \Vdash Fp$. But as t' is the only state where p holds, we must have sRt', and hence s is the intermediate state we were looking for.

Example 18. Suppose we want to model a multi-agent situation, where we have a finite set $S = \{1, \ldots, n\}$ of agents. We let $O = \{\Box_1, \ldots, \Box_n\}$. $\Box_i\varphi$ will now be interpreted as 'Agent i knows φ'. And so we can have formulas $\Box_i(\Box_j\, p \wedge \Box_k q)$ saying 'Agent i knows that Agent j knows p and that Agent k knows q'.

A model for this language is of the form $\mathcal{M} = (W, \{R_1, \ldots, R_n\}, L)$ where R_i is the accessibility relation corresponding to $\Box_i$. This means that the underlying frame is a labeled transition system.

The following example is a good example of how general the new framework is.

Example 19 (Propositional Dynamic Logic)**.** The language of propositional dynamic logic (PDL) has an infinite set of boxes. Each of these boxes has the form $[\pi]$, where π denotes a (non-deterministic) program. The intended meaning of $[\pi]\varphi$ is 'every execution of π from the present state leads to a state bearing the information φ'. The dual assertion $\langle\pi\rangle\varphi$ states that 'some terminating execution of π from the present state leads to a state bearing the information φ'. Now, a very simple idea is going to ensure that PDL is highly expressive: we will make the inductive structure of the programs explicit in PDL's syntax.

Suppose we have fixed some set of basic programs $a, b, c, \ldots$, so that we have the basic modalities $[a], [b], [c], \ldots$ at our disposal. Then we can build more complex programs π over this basis, using the following rules (choice) if π_1 and π_2 are programs, then also $\pi_1 \cup \pi_2$ is a program. The program non-deterministically executes π_1 or π_2.

(composition) if π_1 and π_2 are programs, then also $\pi_1; \pi_2$ is. This program first executes π_1 and then π_2.

(iteration) if π is a program, then so is π^*. This program executes π a finite (possibly zero) number of times.

What this means for the collection of modal operators is that if $[\pi_1]$ and $[\pi_2]$ are boxes, then so are $[\pi_1 \cup \pi_2]$, $[\pi_1; \pi_2]$ and $[\pi_1^*]$. Now, formulas in this language can be used to express properties of program execution. A fairly straightforward example is the formula $\langle \pi^* \rangle \varphi \leftrightarrow \varphi \vee \langle \pi; \pi^* \rangle \varphi$. It says that a state bearing the information φ can be reached by executing π a finite number of times if and only if either we already have the information φ in the current state, or we can execute π once and then find a state bearing the information φ after finitely many more iterations of π. A more complicated formula is

$$\varphi \to ([\pi^*](\varphi \to [\pi]\varphi) \to [\pi^*]\varphi).$$

This is called *Segerberg's axiom* or the *induction axiom.*

Of course nothing stops us from adding more construction rules for programs. Two other such rules are:

(intersection) if π_1 and π_2 are programs, then so is $\pi_1 \cap \pi_2$. This program executes both π_1 and π_2 in parallel.

(test) if φ is a formula, then $\varphi?$ is a program. This program tests whether φ holds, and if so continues; if not, it fails.

A model for PDL has the form $\mathcal{M} = (W, \{R_\pi \mid \pi \text{ is a program}\}, L)$, that is a model is a labeled transition system together with a labelling function. Given the intended meanings of the program constructors we have introduced it is clear that the relations we are interested in are the ones given by the following inductive clauses:

$$\begin{aligned} R_{\pi_1 \cup \pi_2} &= R_{\pi_1} \cup R_{\pi_2} \\ R_{\pi_1; \pi_2} &= R_{\pi_1} \circ R_{\pi_2} \ (= \{(x, y) \mid \exists z (x R_{\pi_1} z \wedge z R_{\pi_2} y)\}) \\ R_{\pi^*} &= (R_\pi)^* \end{aligned}$$

where $(R)^*$ denotes the reflexive and transitive closure of R.

Example 20. PDL can be interpreted on any transition system $(W, R_\pi)_{\pi\in\Pi}$. Such a system is called a *regular* frame if the relations satisfy the conditions of example 19. One can show that a frame $\mathcal{F}$ is regular if and only if $\mathcal{F} \models \Delta\cup\Gamma$ where

$$\Delta := \{p \to ([\pi*](p \to [\pi]p) \to [\pi*]p), \langle\pi*\rangle p \leftrightarrow (p \vee \langle\pi\rangle\langle\pi*\rangle p) \mid \pi \in \Pi\},$$

and

$$\Gamma := \{\langle\pi_1;\pi_2\rangle p \leftrightarrow \langle\pi_1\rangle\langle\pi_2\rangle p, \langle\pi_1 \cup \pi_2\rangle p \leftrightarrow \langle\pi_1\rangle p \vee \langle\pi_2\rangle p \mid \pi_i \in \Pi\}.$$

Exercise 18. Show that regular frames are characterized by the formulas in Δ and Γ, given in example 20

References

[1] Patrick Blackburn, Marteen de Rijke and Yde Venema,
Modal Logic. Cambridge Tracts in Theoretical Computer Science (53) 2004.

[2] Dirk van Dalen,
Logic and Structure. Third edition. Universitext. Springer-Verlag, Berlin, 1994.

[3] Michael R. A. Huth and Mark D. Ryan,
Logic in Computer Science, Modelling and reasoning about systems. Cambridge university press 2000.

[4] Dick de Jongh, Frank Veltman,
Intensional Logics. Course notes,
http://staff.science.uva.nl/~veltman/papers/FVeltman-intlog.pdf

[5] Erik Palmgren,
Konstruktiv Logik, U.U.D.M. Lecture Notes 2002:LN1
http://www.math.uu.se/~palmgren/tillog/klogik02.ps