Manuel Kauers
RISC-Linz

Hagenberg, March 18, 2003

# An Introduction to Modal Logics

## Chapter in Logic II

**Abstract**

This paper gives a first introduction into modal logics. It is based on a lecture of Peter Schmitt at the University of Karlsruhe and is inteded as a starting point for readers who have never touched modal logics before but who are familiar with basic notions of classical propositional and predicate logic. Syntax, semantics, and the underlying intuitive ideas of modal logics are explained and illustrated by examples; some deeper results and directions for further reading are briefly mentioned at the end.

## Contents

# 1 Introduction

The aim of nonclassical logics is to modify the classical logic in such a way that it becomes possible to express things which cannot be expressed otherwise. One approach is to introduce new truth values. Besides of true and false, it might be usefull also to have a truth value "unknown" or to even have infinitely many truth values modeling the truth of a statement by its probability.

A different kind of extending the classical logic is provided by the *modal logics*, which are the topic of this chapter. Modal logics were invented to make it possible also to express that a certain statement $A$ is "possibly" true, or "necessarily," "sometimes," "perhaps," "according to someone's knowledge," etc. This extensions make it possible to reason not only about the absolute truth of a statement, but also about whether someone knows about the truth of a certain statement, for instance.

A typical introductory example of modal logics is the following puzzle, taken — like almost everything in this chapter — from the lecture notes [Schmitt02].

**Example 1** Three philosophers are sitting around wearing black or white hats. They know that at least one of the hats is black. They can see (i.e., they *know*) the colors of the others' hats, and they want to find out the color of their own hat (without cheating, of course).

The first one thinks, and says that he doesn't know the color of his hat. The second one also thinks, but comes to the same conclusion. He says, he doesn't know. But then the third one says that his hat must be black.

Curiously, the third *agent* can extract additional knowledge from the fact that the other agents don't have a certain knowledge. Modal logics allow to formalize the "thoughts" of the three philosophers in the example above and to comprehend their reasoning on a formal level.

The knowledge based interpretation of modal logics obviously popped into the interest of computer science because of its applicability in artificial intelligence. But there are other applications as well. For instance, certain statements may or may not hold in the states of a finite automaton, and one might want to interpret the "necessarily" of a modal logic as "in all reachable states" holds something. This might be usefull to prove the correctness of communication protocols or livenes properties of networks. Other applications are briefly mentioned at the end of this survey.

Modal logics are an old and well studied field, and the magnitude of aspects, applications, and insights does not fit into an article like this, actually not even books can give a somewhat complete overview over modal logics any more. In this chapter, we will just introduce syntax and semantics of modal logics, point out that there are several possible axiom systems, consider some examples, and present some very easy results from *correspondence theory* (to be explained later). At the end, we will briefly mention what else is out there. We based this article on the lecture notes [Schmitt02] which cover the material in considerably more depth, and which also contains pointers to the literature. (They are, however, in German.)

As an English reference, we recommend [GaGuen84] which contains 12 chapters on various aspects of modal logics.

# 2 Syntax and Axiom Systems

## 2.1 The formulas of Modal logics

We will see below that there are several modal logics, not just one. However, they all share the same set of formulas. A formula of modal logic is essentially a formula of propositional logic, extended by two new unary "operators" $\Box$ (box) and $\Diamond$ (diamond). Their interpretation in different circumstances leads to the different modal logics.

**Def. 1** Let $\Sigma$ be a signature for the propositional logic, i.e., a set of boolean variables. By $\mathrm{MFor} = \mathrm{MFor}_\Sigma$, we denote the set of modal logic formulas which is recursively defined as follows.

1. $\Sigma \subseteq \mathrm{MFor}$, $\mathrm{true}, \mathrm{false} \in \mathrm{MFor}$,
2. If $A, B \in \mathrm{MFor}$, then $\neg A$, $A \wedge B$, $A \vee B$, $A \to B$, $A \leftrightarrow B \in \mathrm{MFor}$,
3. If $A \in \mathrm{MFor}$, then $\Box A \in \mathrm{MFor}$ and $\Diamond A \in \mathrm{MFor}$.

Formally, it would be more rigorous to write $(A \vee B)$ instead of $A \vee B$ not to get ambiuguity problems, but we would anyway omit parenteses to make the formulas more human readable. For the logical connectives, we apply the usual precedence rules. For instance, we mean $(A \vee (B \wedge C))$ when we write $A \vee B \wedge C$, and so on. The modal operators are assumed to have highest priority, e.g., we read the formula $\Box A \to A$ as $((\Box A) \to A)$, and likewise for all other logical connectives.

The intuitive meaning of $\Box A$ will depend on the context under consideration. We will interpret $\Diamond A$ always as the "opposite" of $\Box A$ in the sense of $\Diamond A :\leftrightarrow \neg\Box\neg A$. Some possible interpretations of $\Box A$ and the corresponding interpretations of $\Diamond A$ are sumarized in the following table.

| $\Box F$ | $\Diamond F$ |
|---|---|
| $F$ is necessarily true | $F$ is possibly true |
| $F$ is true for ever | $F$ will become true at some point |
| $F$ is true everywhere | $F$ is true somewhere |
| An agent believes $F$ | $F$ is consistent with the things an agent believes |
| An agent knows $F$ | An agent does not know that $F$ is false |
| $F$ is true after each execution of a program | There is an execution of a program after which $F$ is true |
| $F$ is true in all reachable states | There is a reachable state in which $F$ is true. |
| $F$ is obligatory | $F$ is permitted |
| $F$ is true "in all worlds" | There is "a world" where $F$ is true |

In some of the interpretations above, the truth of $\Box A$ depends on parameters, like an agent or a program. If there are several agents or programs involved, it needs to be expressed with respect to which $\Box A$ or $\Diamond A$ is meant. In such cases, we write $\Box_p A$ or $[p]A$ instead of $\Box A$ and $\Diamond_p A$ or $\langle p\rangle A$ instead of $\Diamond A$ to express dependency from, e.g., a program $p$. Strictly speaking, this is then already a *multi modal* logic with more than one box and diamond operator.

**Example 2** Continuing the example of page 2, the situation could be formalized as follows. We call the three philosophers 1, 2, 3. To each of them we associate a box operator expressing knowledge. For instance $\Box_1 A$ expresses that no. 1 knows $A$. Let $B_1, B_2, B_3$ be boolean variables where $B_a$ expresses that $a$'s hat is *black*. The initial knowledge is expressed by $B_1 \vee B_2 \vee B_3$ (there is at least one black hat).

In addition it has to be modelled that each agent sees (and hence knows) the hats of the others. This is expressed by formulas of the form $B_2 \to \Box_3 B_2$, $B_1 \to (\Box_2 B_1 \wedge \Box_3 B_1)$, $(B_1 \wedge B_2) \to \Box_3(B_1 \wedge B_2)$, and so on. Later, we will make use of the formula

$$(\neg B_2 \wedge \neg B_3) \to \Box_1(\neg B_2 \wedge \neg B_3)$$

expressing that no. 1 knows that both the hat of no. 2 and no. 3 are white if this is the case, and of the formula $\neg B_3 \to \Box_2 \neg B_3$ which expresses that no. 2 can see if no. 3 has a white hat.

Finally, the first two agents admit that they don't know the color of their hats: $\neg\Box_1\neg B_1$, $\neg\Box_1 B_1$, $\neg\Box_2\neg B_2$, $\neg\Box_2 B_2$.

The surprising fact, which we want to conclude from the situation just formalized, is that the third agent knows that his hat is black, i.e., we want to prove $\Box_3 B_3$ under the assumptions above.

We cannot prove yet anything, we only have formulas with which we can express facts and claims. Before we can finish the example, we have to agree on axioms that we want to use in the proof. Different interpretations of the box operator lead to different axiom systems, and these different axiom systems define the different, non-equivalent modal logical systems. They are the topic of the next section.

## 2.2 Axiom Systems

Depending on what has to be modeled by the box operator, different axioms may or may not be assumed. Since modal logics are supposed to be extensions of the classical propositional logic, we will however always assume that all propositional tautologies and their modal logic instances are still valid in any modal logic we are interested in. This means in particular, that $\Box A \vee \neg\Box A$ is assumed to hold since it is an instance of $A' \vee \neg A'$ by $A' \leftarrow \Box A$. This does, however, not make any statement about the truth value of $\Box A \vee \Box\neg A$.

In general, we could of corse declare any sentence we wish as axioms of a certain modal logic, but the question at hand is: Which interpretations require which axioms? We consider in the following table several formulas and ask ourselves whether or not these formulas hold in the various interpretations of $\Box F$ which have been stated in the table above. Depending on his intuition, the reader may or may not agree with our decisions.

| $\Box F$ | $F \to \Box F$ | $\Box F \to F$ | $\Box F \to \Diamond F$ | $\Box F \to \Box\Box F$ | $\Box(F \to G) \to (\Box F \to \Box G)$ |
|---|---|---|---|---|---|
| $F$ is necessarily true, it is unavoidable | yes | yes | yes | yes | yes |
| $F$ is true for ever/everywhere | yes | no | no | yes | yes |
| An agent believes $F$ | yes | no | yes | yes | yes |
| An agent knows $F$ | yes | yes | yes | yes | yes |
| $F$ is true after each execution of a program | yes | no | no | no | yes |
| $F$ is obligatory | yes | no | yes | no | yes |
| $F$ is true in all reachable states | yes | no | no | no | yes |

By a *modal logic system*, we understand a set of modal logic axioms, and a set of rules. Different modal logic systems are collected in the following definition, the simplest of which is the system **K**. The names of these systems are given according to a somewhat canonical classification of the various axiom systems and the theories defined by them. This will not be clarified any further in this introduction, but we still want to stick to the "official notion" for consistency.

**Def. 2** A *modal logic system* $S$ is a tuple $S = (A, R)$, where $A \subseteq \mathrm{MFor}$ is a set of *axioms*, and $R \subseteq \mathrm{MFor} \times \mathrm{MFor}$ is a set of *rules*.

By **K**, we denote the system whose axioms are the propositional tautologies together with the additional axiom scheme

$$\Box(A \to B) \to (\Box A \to \Box B) \qquad (A, B \in \mathrm{MFor}).$$

The rules of **K** are *modus ponens* and the so called *generalisation rule*

$$\frac{A,\ A \to B}{B}, \qquad \frac{A}{\Box A}.$$

Other systems are defined by extension of this **K** in the following way:

1. **T** is obtained from **K** by adding the axiom $\Box A \to A$.
2. $\mathbf{S}_4$ is obtained from **T** by adding the axiom $\Box A \to \Box\Box A$.
3. **D** is obtained from **K** by adding the axiom $\Box A \to \Diamond A$.
4. $\mathbf{S}_{4.2}$ is obtained from $\mathbf{S}_4$ by adding the axiom $\Diamond\Box A \to \Box\Diamond A$.

This list could be extended ad infinitum.

The definition of logical consequence, which allows us to do proofs, carries over literally from the according definition in classical logic. It depends, however, on the underlying modal logic system.

**Def. 3** Let $F, G \in \mathrm{MFor}$ be modal logic formulas, and let **S** be a modal logic system. We say $G$ to be a *modal logical consequence* of $F$ in **S**, written $F \vdash^{\mathbf{S}} G$, if there is a *proof* of $G$ from $F$. Such a proof is a sequence $F_0, \ldots, F_m$ in MFor where $F_m = G$, and each $F_i$ is either equal to $F$, an axiom of **S**, or it is obtained by application of a rule of **S** to formulas that appear earlier in the sequence.

### 2.3 Examples

We are now equipped with the necessary tools that are needed to do *proofs* in a modal logic. Actually, there is nothing special about proving in a modal logic: It is just starting with axioms and applying rules. This will be illustrated in some easy examples, and in a continuation of the example with the three philosophers.

**Example 3** We assume that the underlying modal logic is $\mathbf{K}$, and prove two simple sentences that are valid in $\mathbf{K}$.

- Claim: $(\Box A \wedge \Box B) \leftrightarrow \Box(A \wedge B)$.
  Proof: "$\leftarrow$" Assume $\Box(A \wedge B)$ and prove $\Box A \wedge \Box B$.

| | |
|---|---|
| 1 (Ax) | $\Box((A \wedge B) \to A) \to (\Box(A \wedge B) \to \Box A)$ |
| 2 (Ax) | $(A \wedge B) \to A$ |
| 3 (2) | $\Box((A \wedge B) \to A)$ |
| 4 (1, 3) | $\Box(A \wedge B) \to \Box A$ |
| 5 (Ax) | $\Box((A \wedge B) \to B) \to (\Box(A \wedge B) \to \Box B)$ |
| 6 (Ax) | $(A \wedge B) \to B$ |
| 7 (6) | $\Box((A \wedge B) \to B)$ |
| 8 (5, 7) | $\Box(A \wedge B) \to \Box B$ |
| 9 (4, 8) | $\Box(A \wedge B) \to (\Box A \wedge \Box B)$ |
| 10 (As) | $\Box(A \wedge B)$ |
| 11 (9, 10) | $\Box A \wedge \Box B$. |

  "$\to$" Assume $\Box A \wedge \Box B$ and prove $\Box(A \wedge B)$.

| | |
|---|---|
| 12 (As) | $\Box A \wedge \Box B$ |
| 13 (Ax) | $\Box(B \to (A \to (A \wedge B))) \to (\Box B \to \Box(A \to (A \wedge B))$ |
| 14 (Ax) | $B \to (A \to (A \wedge B))$ |
| 15 (14) | $\Box(B \to (A \to (A \wedge B)))$ |
| 16 (15, 13) | $\Box B \to \Box(A \to (A \wedge B))$ |
| 17 (Ax) | $\Box(A \to (A \wedge B)) \to (\Box A \to \Box(A \wedge B))$ |
| 18 (16, 17) | $\Box B \to (\Box A \to \Box(A \wedge B))$ |
| 19 (18) | $(\Box A \wedge \Box B) \to \Box(A \wedge B)$ |
| 20 (12, 19) | $\Box(A \wedge B)$. |

- Claim: $(\Box A \vee \Box B) \to \Box(A \vee B)$.
  Proof:

| | |
|---|---|
| 1 (Ax) | $\Box(A \to (A \vee B)) \to (\Box A \to \Box(A \vee B))$ |
| 2 (Ax) | $A \to (A \vee B)$ |
| 3 (2) | $\Box(A \to (A \vee B))$ |
| 4 (1, 3) | $\Box A \to \Box(A \vee B)$ |
| 5 (Ax) | $\Box(B \to (A \vee B)) \to (\Box B \to \Box(A \vee B))$ |
| 6 (Ax) | $B \to (A \vee B)$ |
| 7 (6) | $\Box(B \to (A \vee B))$ |

| | |
|---|---|
| 8 (5, 7) | $\Box B \to \Box(A \vee B)$ |
| 9 (4, 8) | $(\Box A \vee \Box B) \to \Box(A \vee B)$ |

A natural question is to ask for the converse of the last example. Is $\Box(A \vee B) \to (\Box A \vee \Box B)$ also valid in **K**? No, it is not, but we cannot prove it at this point. In predicate logic, to prove that a certain formula $F$ is not valid, it suffices to construct an interpretation (i.e. a universe with concrete functions and predicates for every function and predicate symbol of the signature) where $F$ does not hold. We will define interpretations for modal logics in the next section, and we will give a counterexample to $\Box(A \vee B) \to (\Box A \vee \Box B)$.

Before we turn to the formal interpretations of modal logics, we first finish the example with the philosophers and their hats.

**Example 4** We take the situation of the three philosophers as it was modeled in Example 2 and formaly deduce that the third philosopher indeed knows that his hat is black. By F, we refer to a fact, i.e., an axiom of **K** or a formula which models the situation. Instances of the axiom $\Box(A \to B) \to (\Box A \to \Box B)$ together with modus ponens will also be used as rule.

| | |
|---|---|
| 1 (F) | $B_1 \vee B_2 \vee B_3$ |
| 2 (1) | $(\neg B_2 \wedge \neg B_3) \to B_1$ |
| 3 (2) | $\Box_1((\neg B_2 \wedge \neg B_3) \to B_1)$ |
| 4 (3) | $\Box_1(\neg B_2 \wedge \neg B_3) \to \Box_1 B_1$ |
| 5 (F) | $\neg\Box_1 B_1$ |
| 6 (4, 5) | $\neg\Box_1(\neg B_2 \wedge \neg B_3)$ |
| 7 (F) | $(\neg B_2 \wedge \neg B_3) \to \Box_1(\neg B_2 \wedge \neg B_3)$ |
| 8 (6, 7) | $\neg(\neg B_2 \wedge \neg B_3)$ |
| 9 (8) | $B_2 \vee B_3$ |
| 10 (9) | $\neg B_3 \to B_2$ |
| 11 (10) | $\Box_2(\neg B_3 \to B_2)$ |
| 12 (11) | $\Box_2\neg B_3 \to \Box_2 B_2$ |
| 13 (F) | $\neg\Box_2 B_2$ |
| 14 (12, 13) | $\neg\Box_2\neg B_3$ |
| 15 (F) | $\neg B_3 \to \Box_2\neg B_3$ |
| 16 (14, 15) | $B_3$ |
| 17 (16) | $\Box_3 B_3$. |

Informally speaking, the argument is as follows. The first philosopher does not know the color of his hat. But he knows that there is at least one black hat. If both the others did not have black hats, he could see this and conclud that his hat is black. From his statement that he doesn't know the others can conclude that one of their hats must be black. If the third's hat were white, the second's would have to be black, and he would know about that. But since he says that he doesn't know either, the only remaining possibility is that no. 3 must have a black hat. Note that the third philosopher would even come to this conclusion if he were blind.

# 3 Semantics: Kripke Structures

## 3.1 Kripke Structures

The structures in which modal logical formulas are interpreted are called *Kripke Structures.*

**Def. 4** A *Kripke Frame* is a pair $\mathcal{F} = (G, R)$ where $G$ is a set and $R$ is a relation on $G$. The elements $g \in G$ are called the *worlds* of $\mathcal{F}$, and $R$ is called the *reachability relation* of worlds.

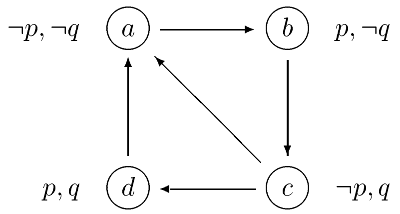
**Def. 5** A *Kripke Structure* to a signature $\Sigma$ is a triple $\mathcal{K} = (G, R, v)$ where $(G, R)$ is a *Kripke Frame*, and $v \colon G \times \Sigma \to \{0, 1\}$ is a function which assigns truth values to each boolean variable.

Kripke frames where $G$ is finite can be drawn as directed graphs. The worlds of the structure correspond to the nodes of the graph, and $R$ corresponds to its edges. The interpretation function $v$ acts in each individual world like a usual interpretation function of propositional logic.

**Example 5** Let $\Sigma = \{p, q\}$ be a signature and $\mathcal{K} = (G, R, v)$ be the Kripke structure to $\Sigma$ defined by $G = \{a, b, c, d\}$, $R = \{(a, b), (b, c), (c, d), (d, a)\}$ and

$$\begin{array}{cc} v(a,p) = 0 & v(a,q) = 0 \\ v(b,p) = 0 & v(b,q) = 1 \\ v(c,p) = 1 & v(c,q) = 0 \\ v(d,p) = 1 & v(d,q) = 1 \end{array}$$

The structure $\mathcal{K}$ looks as follows.




For a world $g \in G$ and a variable $s \in \Sigma$, the annotation $s$ means $v(g, s) = 1$, and $\neg s$ is to be read as $v(g, s) = 0$.

The next step is the definition of *validity* of a modal logic formula in a certain world of a Kripke structure, or in an entire Kripke structure.

## 3.2 Interpretation

First we define what is means that a certain formula $F$ is valid in a certain world of a Kripke structure. The propositional connectives behave as expected, there is no difference to the usual interpretation in propositional logic. Different worlds do not "influence each other" with respect to propositional logic.

The difference comes only into play in the interpretation of the modal logic operators $\Box$ and $\Diamond$. These formulas relate the truth of a world $g$ to the truth in the worlds to which $g$ is related by means of the Kripke structure's relation.

The concrete definition is as follows:

**Def. 6** Let $\mathcal{K} = (G, R, v)$ be a Kripke structure, $g \in G$, and $A, B \in$ MFor. We write $(\mathcal{K}, g) \models A$ (also $\mathcal{K}, g \models A$, or even $g \models A$ if $\mathcal{K}$ is clear) to say that $A$ is *valid* in the world $g$ of the Kripke structure $\mathcal{K}$.

The validness of formulas if recursively defined as follows, where $A, B \in$ MFor.

1. $g \models p$ for $p \in \Sigma$ iff $v(g, p) = 1$,
2. $g \models \neg A$ iff not $g \models A$,
3. $g \models A \vee B$ iff $g \models A$ or $g \models B$,
4. $g \models \Box A$ iff for all $h \in G$ with $R(g, h)$ we have $h \models A$,
5. $g \models \Diamond A$ iff there is an $h \in G$ with $R(g, h)$ such that $h \models A$.

If $\wedge, \rightarrow, \leftrightarrow$ are taken as part of the language rather than as abbreviations, their semantics is defined in the obvious way, analogous to $g \models A \vee B$.

If a formula $A \in$ MFor is valid in all worlds $g$ of a Kripke structure $\mathcal{K} = (G, R, v)$, then we say $A$ to be *globally valid* in $\mathcal{K}$, writing $\mathcal{K} \models A$. In this case, we call $\mathcal{K}$ a *model* of $A$.

**Example 6** In the setting of Example 5, we have for instance

$$\begin{array}{ll} \mathcal{K}, d \models p \wedge q, & \mathcal{K}, c \models \Box(p \leftrightarrow q), \\ \mathcal{K}, c \not\models \Box p & \mathcal{K}, c \models \Diamond p, \\ \mathcal{K}, c \models \Diamond\neg p, & \mathcal{K}, b \models p \wedge \Diamond q, \end{array}$$

and so on. In this particular example, we can also make the general statement $\mathcal{K}, g \models \Box F \rightarrow \Diamond F$ for every $F \in$ MFor and every $g \in \{a, b, c, d\}$. (Is this true in every Kripke structure?) We conclude that $\mathcal{K}$ is in particular a model of $\Box p \leftrightarrow \Diamond p$.

In propositional logic, $F$ is said to be a *consequence* of $G$ if every model of $F$ is also a model of $G$. There are two generalisations of this notion in modal logic, they are introduced in the following definition.

**Def. 7** Let $A, B \in$ MFor.

1. $B$ is said to be a *global modal logical consequence* of $A$ if for every Kripke structure $\mathcal{K} = (G, R, v)$ we have $B$ is true in all worlds of $G$ provided that $A$ is true for all worlds of $G$.
2. $B$ is said to be a *local modal logical consequence* of $A$ if for every Kripke structure $\mathcal{K} = (G, R, v)$ and every world $g$, $(\mathcal{K}, g) \models A$ implies that $(\mathcal{K}, g) \models B$.

We write $A \models_G B$ in the first case and $A \models_L B$ in the second situation.

## 3.3 Properties of Kripke Structures

The notion of Kripke structures formalises the interpretation of $\Box A$ which says that "$A$ is true in all reachable states." It is rather hard to recognize other possible interpretations like "An agent knows $A$" in a Kripke structure. However, there is one argument why Kripke structures are taken as interpretations in general: They turn out to be consistent with the easiest modal logical system which is of interest, namely the system **K** as defined in Def. 2.

**Theorem 1** Let $\mathcal{K} = (G, R, v)$ be a Kripke structure and $A, B \in \mathrm{MFor}$. Then

1. $\mathcal{K} \models A$ if $A$ is a tautology,
2. If $\mathcal{K} \models A$ and $\mathcal{K} \models A \to B$, then $\mathcal{K} \models B$.
3. If $\mathcal{K} \models A$, then $\mathcal{K} \models \Box A$,
4. $\mathcal{K} \models \Box(A \to B) \to (\Box A \to \Box B)$,

**Proof.** Part 1. and 2. are immediate from the definition of semantics and left to the reader.

3. $\mathcal{K} \models A$ implies $\mathcal{K} \models \Box A$. Assume that $\mathcal{K} \models A$. This means by definition that $\mathcal{K}, g \models A$ for each $g \in \mathcal{K}$. In order to show $\mathcal{K} \models \Box A$ we have to show that $\mathcal{K}, g \models \Box A$ for each $g \in G$. Fix a $g \in G$. To show $\mathcal{K}, g \models \Box A$, we have to show that $\mathcal{K}, h \models A$ for each $h \in G$ with $R(g, h)$. By assumption, we have $\mathcal{K}, h \models A$ for every $h \in G$, thus also for the ones with $R(g, h)$.

4. $\mathcal{K} \models \Box(A \to B) \to (\Box A \to \Box B)$. Take a world $g \in G$, assume $\mathcal{K}, g \models \Box(A \to B)$ and $\mathcal{K}, g \models \Box A$, and show $\mathcal{K}, g \models \Box B$. Let $h \in G$ be any world which is reachable from $g$, i.e., $R(g, h)$. Then $\mathcal{K}, g \models \Box(A \to B)$ yields $\mathcal{K}, h \models A \to B$, $\mathcal{K}, g \models \Box A$ yields $\mathcal{K}, h \models A$, and to show $\mathcal{K}, g \models \Box B$, it suffices to show $\mathcal{K}, h \models B$ since $h$ is arbitrary:

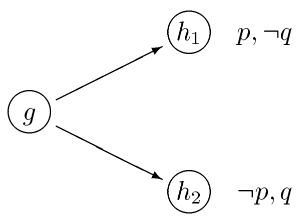
$$\Box(A \to B), \Box A \quad g \longrightarrow h \quad A \to B, A$$

By Modus Ponens, which works due to 2., we can conclude that $\mathcal{K}, h \models B$, and hence we have $\mathcal{K}, g \models \Box B$ since $h$ was arbitrary. ∎

Now we can return to the question that we could not answer in Section 2.3 on page 6: Does

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

hold in the system **K**? We already said that this is not the case, and we can prove it now by simply constructing a Kripke structure which has a world in which the left hand side is true but the right hand side is not. This is not difficult to do, here is a possible solution:



Consider the world $g$ in this structure. For this world we have $g \models \Box(p \vee q)$ because in every reachable world, one of the two variables is true. On the other hand, none of the variables is true in all reachable worlds, i.e., $g \not\models \Box p \vee \Box q$.

Kripke structures provide a semantics to the modal logical system **K**. But how should the semantics of other systems like **T**, **D**, or $\mathbf{S}_4$ be understood? As these systems are defined as refinements of **K**, they are obtained by adding further axioms to **K**. Still, Kripke structures may serve to define the semantic part of these systems, only the additional axioms restrict the class of possible underlying Kripke frames. As an example, we consider the system **T** which has, according to Def. 2, the additional axioms $\Box A \to A$ ($A \in \mathrm{MFor}$).

**Theorem 2** Let $(G, R)$ be a *reflexive* Kripke frame, i.e., every world $g \in G$ is in relation with itself, $R(g, g)$ holds. Then in every Kripke structure $\mathcal{K} = (G, R, v)$ having this frame every formula $\Box A \to A$ ($A \in \mathrm{MFor}$) holds.

**Proof.** Let $g \in G$ and assume that $g \models \Box A$ for some $A \in \mathrm{MFor}$ fixed. Then, by the definition of $\Box$, for every $h \in G$ with $R(g, h)$ we have $h \models A$. But $R$ is reflexive by assumption, so $g$ itself is such an $h$ and we have $g \models A$ which was to prove. ∎

Conversely, if $(G, R)$ is a frame such that $\Box A \to A$ holds in every structure $(G, R, v)$, then $(G, R)$ must be reflexive. Otherwise, $v$ could be easily constructed in such a way that it fulfills $\Box A$ but violates $A$, contradicting the claim. (The reader might try this as an exercise.) We observe that there is obviously some kind of "equivalence" between the formula $\Box A \to A$ of modal logic and the predicate logic formula $\forall x\ R(x, x)$. This equivalence will be studied to some more extend in the next section.

Similar considerations are possible for other modal logic systems. The formula $\Box A \to \Box\Box A$ which is the additional axiom in $\mathbf{S}_4$, restricts the Kripke frames to those whose relation is transitive, together with $A \to \Box A$ this means that a formula $F$ is tautologic in $\mathbf{S}_4$ if and only if it is true in all Kripke structures whose underlying frames are reflexive and transitive.

For the system **D**, the corresponding class of Kripke frames is that where the relation is such that every world has at least one world to which it is related. The axiom $\Diamond\Box A \to \Box\Diamond A$ of $\mathbf{S}_{4.2}$ restricts the choice for the relation of a Kripke frame to the class of relation that fulfills $\exists x\ \forall y\ R(x, y) \to \forall x\ \exists y\ R(x, y)$.

# 4 Correspondence Theory

The preceeding two sections introduced syntax and semantics of modal logics and tried to give some intuition about what modal logics are dealing with. But this is actually only a starting point for the study of many aspects of this kind of logics each of which can be studied quite deeply. Of the many possible questions we will give a rather short overview over the *correspondence theory*, because we think that the questions treated here naturally arise in the way we introduced syntax and semantics of modal logics before. Other aspects and applications will be dealt with not at all in this survey, but some of them will at least be mentioned in Section 5 to give at least an idea about what else is going on in the field.

## 4.1 Relating Kripke Structures to Formulas

We have seen in Section 3.3 already that formulas of modal logic can be viewed as selectors that restrict the class of all Kripke frames to the class where these formulas are valid independend of the choice of the interpretation function $v$. For instance, we have seen that $\Box A \to A$ is valid for all possible interpretation functions $v$ if and only if the underlying frame is reflexive. This correspondence is generalized in the following definition.

**Def. 8** Let $C$ be a class of Kripke frames and $F \in \text{MFor}$. We say that $C$ is *characterized* by $F$ iff for all Kripke frames $(G, R)$ we have that $(G, R) \in C$ if and only if $(G, R, v) \models F$ for every $v$.

In this sense, $\Box p \to p$ characterises the class of all reflexive Kripke frames. Some more examples are sumarized in the following table. The left column shows a modal logical formula, and the right column contains a formula from predicate logic describing the characterized Kripke frame, or, more precisely, its relation $R$.

| | |
|---|---|
| $\Box p \to p$ | $\forall x\ R(x, x)$ |
| $p \to \Box\Diamond p$ | $\forall x\ \forall y\ (R(x, y) \to R(y, x))$ |
| $\Box p \to \Diamond p$ | $\forall x\ \exists y\ R(x, y)$ |
| $\Box p \to \Box\Box p$ | $\forall x\ \forall y\ \forall z\ (R(x, y) \wedge R(y, z) \to R(x, z))$ |
| $\Diamond p \to \Box\Diamond p$ | $\forall x\ \forall y\ \forall z\ (R(x, y) \wedge R(x, z) \to R(y, z))$ |
| $\Box\Box p \to \Box p$ | $\forall x\ \forall y\ (R(x, y) \to \exists z\ (R(x, z) \wedge R(z, y)))$ |

Correspondence theory studies the relationship between modal logic formulas and the Kripke frames they characterize. Besides of extending the above list ad infinitum, this subfield aims to find more general results as well. Two of these more interesting results will be presented in the sequel. They are concerned with the comparison between the expressive power of modal logic and predicate logic in the sense of the following two main questions:

1. Given a formula $F'$ of first order predicate logic, does there always exist a formula $F \in \text{MFor}$ such that $F$ characterizes the class of frames defined by $F'$?

2. Given a formula $F \in \text{MFor}$, does there always exist a formula $F'$ of first order predicate logic which defines the class of frames characterized by $F$?

Surprisingly enough, the answer to both of these questions is negative; we will see counterexamples.

## 4.2 Irreflexive Frames are not characterisable

To see that not every kind of relation which can be described by first order logic can also be characterized by a modal logic formula, we consider *irreflexive* frames. A frame $(G, R)$ is called irreflexive if no world is related to itself, $\forall x\ \neg R(x, x)$.

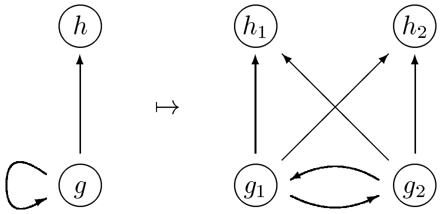
We will proceed by defining a procedure which transforms every Kripke structure $\mathcal{K}$ into an irreflexive structure $\mathcal{K}^*$ in such a way that $\mathcal{K} \models F \iff \mathcal{K}^* \models F$

for all $F \in$ MFor. The assumption that $F$ be a formula that characterizes irreflexivity then leads to a contradiction in the following way. Consider a structure $\mathcal{K} = (G, R, v)$ where $G = \{g\}$ and $R(g, g)$. Then $\mathcal{K}^*$ is irreflexive by construction and by assumption $\mathcal{K}^* \models F$. But then also $\mathcal{K} \models F$ although $\mathcal{K}$ is not irreflexive. This cannot be true if $F$ characterizes irreflexivity, and hence such an $F$ cannot exist.

We define $\mathcal{K}^*$ in the following way. To a given structure $\mathcal{K} = (G, R, v)$ put $\mathcal{K}^* = (G^*, R^*, v^*)$ such that

$$\begin{aligned} G^* &:= \{(g,0),(g,1) : g \in G\}, \\ R((g,i),(h,j)) &\iff R(g,h) \text{ and } g \neq h \\ R((g,i),(g,j)) &\iff R(g,g) \text{ and } i \neq j \\ v^*((g,i),p) &:= v(g,p). \end{aligned}$$

Intuitively, $\mathcal{K}^*$ is obtained from $\mathcal{K}$ by creating two copies of $\mathcal{K}$, relating each world $g$ to $h$ and its copy $h'$ if it was originally related to $R(g, h)$. The original relation ships are kept unless they are self references. It is obvious that $\mathcal{K}^*$ is not reflexive, no matter which structure the original $\mathcal{K}$ has had.




It remains to show that going from $\mathcal{K}$ to $\mathcal{K}^*$ does not change the truth of the any modal logic formula. This is the content of the following lemma.

**Theorem 3** Let $\mathcal{K} = (G, R, v)$ be a Kripke structure. Then, for every $F \in$ MFor and every world $g \in G$ we have

$$\mathcal{K}, g \models F \iff \mathcal{K}^*, (g,0) \models F \iff \mathcal{K}^*, (g,1) \models F.$$

As a consequence, $\mathcal{K} \models F \iff \mathcal{K}^* \models F$.

**Proof.** We prove two the claims $\mathcal{K}, g \models F \iff \mathcal{K}^*, (g,0) \models F$ and $\mathcal{K}, g \models F \iff \mathcal{K}^*, (g,1) \models F$ by simultanous induction on the structure of $F$.

We only consider the case $\Box F$. The base case follows immediately from the definition of $v^*$, the boolean connectives are easy, and the case $\Diamond F$ proceeds by an argument that is analogous to what follows. So assume $\mathcal{K}, g \models F \iff \mathcal{K}^*, (g,i) \models F$ as induction hypothesis and show that these equivalences do also hold with $\Box F$ in place of $F$.

"$\Rightarrow$" Assume $\mathcal{K}, g \models \Box F$ and show $\mathcal{K}^*, (g,i) \models \Box F$. Let $u = (h,j) \in G^*$ be arbitrary with $R^*((g,i),u)$. We have to show $\mathcal{K}^*, u \models F$.

Case 1: $h = g$, $i \neq j$. By definition of $R^*$, we have that $R^*((g,i),(g,j))$ implies $R(g,g)$. Together with the assumption $\mathcal{K}, g \models \Box F$ this gives $\mathcal{K}, g \models F$ and by induction hypothesis this implies $\mathcal{K}^*, (g,j) \models F$.

Case 2: $h \neq g$. By definition of $R^*$, we have that $R^*((g,i),(h,j))$ implies $R(g,h)$. Together with the assumption $\mathcal{K}, g \models \Box F$ this implies $\mathcal{K}, h \models F$ and by induction hypothesis this implies $\mathcal{K}^*, (h,j) \models F$.

"$\Leftarrow$" As for the converse direction, assume $\mathcal{K}^*, (g, i) \models \Box F$ and show $\mathcal{K}, g \models \Box F$. Let $h \in G$ be such that $R(g, h)$. We have to show $\mathcal{K}, h \models F$. By construction of $R$ we have that $R((g, i), (h, 1 - i))$, independent of whether or not $g = h$. Together with the assumption $\mathcal{K}^*, (g, i) \models \Box F$ this leads to $\mathcal{K}^*, (h, 1 - i) \models F$ the claim $\mathcal{K}, h \models F$ follows again by induction hypothesis. ■

This completes also the proof that irreflexive frames are not characterizable, using the argument that has been carried out in the beginning. Generally speaking, this shows that not every class of frames which can be characterized by a first order formula can also be characterized by a formula of modal logics.

## 4.3 Being noetherian is characterisable

We now show an example of a class of frames which is characterizable but which nevertheless can not be defined by any formula of first order predicate logic. It will turn out that the class of transitive noetherian frames serves our purpose. First we recall the definition of noetherian.

**Def. 9** A relation $R$ on a set $G$ is called *noetherian*, if there is no infinite sequence $g_1, g_2, g_3, \ldots$ with $R(g_1, g_2)$, $R(g_2, g_3), \ldots$. This means that there is no infinite ascending sequence in $G$ wrt. $R$.

It is an easy corollary to the compactness theorem that it can not be formalized by a formula $\Phi$ of first order logic that a binary predicate symbol $R$ be noetherian and transitive. On the other hand, the following theorem shows that the class of noetherian transitive frames in characterizable by a formula of modal logic.

**Theorem 4** The formula $\Box(\Box p \to p) \to \Box p$, known as the *Lö b axiom*, characterizes the class of all transitive Kripke frames whose relation $R$ is noetherian. This means that there is no infinite chain $g_0, g_1, g_2 \ldots$ of worlds such that $R(g_0, g_1)$, $R(g_1, g_2)$, $R(g_2, g_3), \ldots$.

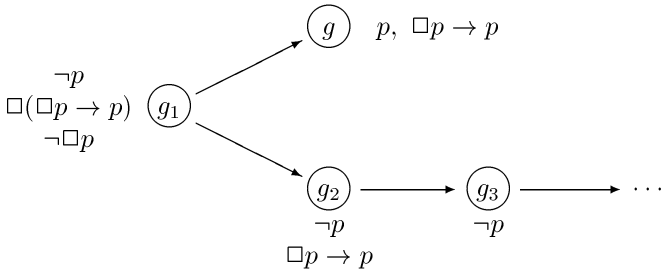
**Proof.** We have to show that the Lö b axiom is satisfied in a frame $\mathcal{F}$ for any interpretation if and only if $\mathcal{K}$ is transitive and noetherian.

"$\Rightarrow$" Let $\mathcal{F} = (G, R)$ be a Kripke frame where the Lö b axiom is valid. We have to show that $\mathcal{F}$ is transitive and noetherian.

Assume it is not transitive. Then there are worlds $g_1, g_2, g_3$ with $R(g_1, g_2)$, $R(g_2, g_3)$ but not $R(g_1, g_3)$. Put $g_1 \models p$, $g_2 \models \neg p$ and $g_3 \models \neg g$. Then $g_1 \models \Box(\Box p \to p)$ but $g_1 \not\models \Box p$ in contradiction to the axiom.

$$\begin{array}{ccccc} g_1 & \longrightarrow & g_2 & \longrightarrow & g_3 \\ p & & \neg p & & \neg p \\ \Box(\Box p \to p) & & \Box p \to p & & \\ \neg\Box p & & & & \end{array}$$

Now assume that $\mathcal{F}$ is not noetherian, and let $g_1, g_2, g_3, \ldots$ be an infinite ascending chain. Put $g_i \models \neg p$ for all $g_i$ in the chain and $g \models p$ for all other worlds.




It is clear that $g_1 \models \neg\Box p$ because $g_2 \models \neg p$. But we will show that $g_1 \models \Box(\Box p \to p)$ then we have a contradiction to Lö b's axiom in $g_1$. To show $g_1 \models \Box(\Box p \to p)$, we have to show $h \models \Box p \to p$ for any $h$ in relation to $g_1$.

Case 1: $h$ does not belong to the infinite chain. Then we have $h \models p$ and therefore also $h \models \Box p \to p$ independent of whether or not $h \models \Box p$.

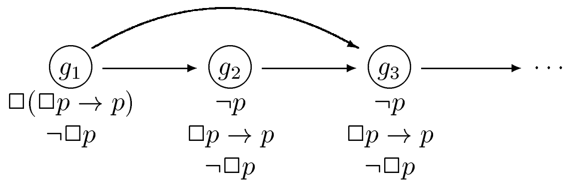
Case 2: $h$ is a member of the infinite chain, wlog. $h = g_2$. Then $h \models \neg p$, but still $h \models \Box p \to p$ because $h$ has a successor $h'$ with $h' \models p$ witnessing $h \not\models \Box p$.

"$\Leftarrow$" As for the converse direction, we assume that $\mathcal{F}$ is transitive and noetherian and conclude that Lö b's axiom is valid. This amounts to the same as showing that if the Lö b axiom is not valid in a transitive frame, then this frame is not noetherian.

Let $\mathcal{F} = (G, R)$ be a transitive frame where Lö b's axiom is valid. This means that there is a world $g_1 \in G$ such that

$$g_1 \not\models \Box(\Box p \to p) \to \Box p,$$

i.e., $g_1 \models \Box(\Box p \to p)$ but $g_1 \models \Diamond\neg p$. The latter implies the existence of an $g_2$ with $R(g_1, g_2)$ and $g_2 \models \neg p$. On the other hand $g_1 \models \Box(\Box p \to p)$ implies also $g_2 \models \Box p \to p$. Both statements about $g_2$ rule out the possibility that $g_2 \models \Box p$, and hence there must be a $g_3 \in G$ with $g_3 \models \neg p$. By transitivity, $g_1 \models \Box(\Box p \to p)$ also implies $g_3 \models \Box p \to p$, and so on.




An infinite chain $g_1, g_2, g_3 \ldots$ with $R(g_i, g_{i+1})$ is constructed by repeating this argument, showing that $(G, R)$ is not noetherian, as desired. ■

This shows in particular that not every class of frames which can be characterized by a modal logic formula can also be characterized by a first order formula.

However, it is always possible to give a formula of second order predicate logic that describes any arbitrary class of frames that can be characterized by a formula in MFor.

# 5 Other Aspects and Applications

In this concluding section, we briefly mention other topics of modal logics, somehow arbitrarily choosing aspects which might be of interest. We only intend to put a spotlight on each of these subfields, roughly giving an outline of the underlying ideas but certainly refusing correctness, completeness, or formal rigorosity. The interested reader is refered to [Schmitt02] who covers all mentioned aspects in some more depth and gives references for further reading.

## 5.1 Completeness Results

One of the main results of classical predicate logic is that the relations $\vdash$ and $\models$ are equivalent, for each two formulas $F, G$ of first order logic, we have $F \vdash G$ if and only if $F \models G$. This fact is known as the *completenes* of $\vdash$.

In the first sections, we also defined $\vdash$ and $\models$ for modal logics and it is natural to ask for their relationship. We only mention some results in this direction.

**Theorem 5** Let $A \in \mathrm{MFor}$. The modal logical consequence $\vdash^{\mathbf{S}}$ of a modal logical system $\mathbf{S}$ was defined in Def. 3. Then

1. $\vdash^{\mathbf{K}} A$ if and only if $\mathcal{K} \models A$ for every Kripke structure $\mathcal{K}$.
2. $\vdash^{\mathbf{T}} A$ if and only if $\mathcal{K} \models A$ for every reflexive Kripke structure $\mathcal{K}$.
3. $\vdash^{\mathbf{S}_4} A$ if and only if $\mathcal{K} \models A$ for every reflexive transitive Kripke structure $\mathcal{K}$.

For the direction from left to right it is sufficient to check that all allowed Kripke frames fullfill the axioms of the system under consideration. For the theories $\mathbf{K}$ and $\mathbf{T}$, we did this in Section 3.3, the arguments for other theories works similar.

To prove the converse is however no longer that cheap. The idea is to define *canonical Kripke frames* which are in a certain sense "as croase as possible." Then it is shown that it suffices to show the claim for these canonical models, and for them the claim can be checked more or less by inspection.

## 5.2 Decidability of Modal Logics

The main advantage of modal logics over the entire predicate logic is the fact that it can be algorithmically decided if a formula $F \in \mathrm{MFor}$ holds in a modal logical system or not.

To each formula $A \in \mathrm{MFor}$ it is possible to compute a bound $M \in \mathbb{N}$ such that if there is any Kripke structure $\mathcal{K}$ with $\mathcal{K} \models A$ then there is also a Kripke structure $\mathcal{K}'$ with less than $M$ worlds and $\mathcal{K}' \models M$. To each $M$, there are only finitely many possible Kripke structures with less than $M$ worlds. They can be enumerated and the validnes of $A$ can be checked.

Along this outline it can be established that modal logics are decidable, but this proof does obviously not yield to an algorithm of practical relevance. For practical purposes, there are several possibilities to adapt the ideas of classical calculi like the tableau method also to modal logics. This gives complete and correct proving procedures which under appropriate fairnes conditions become decision procedures.

## 5.3 Modal Predicate Logic

Up to now, we only dealt with propositional modal logics. What happens if we admit also quantifiers and predicates involving terms and variables within modal logical formulas? Of course, this is possible. The notion of a Kripke structure can be extended to modal predicate logic by saying that each world be a usual universe for predicate logic with its own interpretation for function and predicate symbols. Box and diamond operator relate worlds to each other as they do in propositional modal logics.

In principle, this all is possible to do, but there is one reason why modal predicate logics are not studied as extensively as propositional modal logics are: They are simply not interesting. By extending a predicate signature $\Sigma$ by two new predicate symbols $w(x)$ and $r(x,y)$ it is rather easy to construct a procedure that transforms a modal predicate logic formula $A$ over $\Sigma$ to a pure predicate logic formula $A'$ in the extended signature and which transforms a predicate Kripke structure to $\Sigma$ into a usual predicate structure $K'$ over the extended signature in such a way that

$$K \models A \iff K' \models A'.$$

This means that box and diamond operator do not add anything new to predicate logic, and hence they are rather useless.

## 5.4 Specification Languages

In the contexts of knowledge representation and artificial intelligence, as well as in software engeneering, there is a need of formal languages which allow to express properties of certain concepts and relations between these concepts. As an example, one might think of an object oriented design including classes Window and Menu where the assertion has to be expressed that each window object has a menu object associated to it.

Many specification languages have been developed for purposes like this, and most of them can up to syntax be viewed as a restricted subset of predicate logic. Some of them are even restricted enough to be viewed as modal logics. This can under appropriate circumstances allow to prove assertions in a language much easier than in predicate logic, making use of the fact that modal logics are decidable.

## 5.5 Dynamic Logic

Another appliation of modal logics is *program verification*. The interpretation of $\Box A$ as “after each execution of a certain program does $A$ hold” has already been one of the introducing examples. Dynamic logic is based on this interpretation and extends the syntax of modal logic to a multimodal logic with a box operator for each individual program of the programing language. It allows to specify assertions about programs, formalizing, for instance, the correctnes of a program, and it also gives rise to a calculus for verifying such assertions.

Programs are composed recursively out of assignments $x := t$ ($x$ a variable, $t$ a term), of loops $\underline{\textbf{while}}$ $A$ $\underline{\textbf{do}}$ $P$ $\underline{\textbf{end}}$ ($A$ a formula and $P$ a program), and of branches $\underline{\textbf{if}}$ $A$ $\underline{\textbf{then}}$ $P_1$ $\underline{\textbf{else}}$ $P_2$ $\underline{\textbf{end}}$ ($A$ a formula, $P_1, P_2$ programs). Formulas are formulas of multimodal predicate logic with a box operator $[p]$ and a diamond operator $\langle p \rangle$ for each program $p$.

A bit of work is needed to set up the semantics of dynamic logic properly, but the intuitive meaning of formulas is as indicated in Section 2: A formula $[p]F$ is meant to be true if and only if $F$ is true after each execution of the program $p$. For instance, $[x := 1](x = 1)$ and $(f(a) = 7) \rightarrow [x := a](f(x) = 7)$ are true.

We claimed in 5.3 that modal predicate logics are not interesting because they can be embedded into classical first order logic at the price of two additional predicate symbols. However, dynamic logic is multi modal with infinitely many box and diamond operators — one for every program. To embed this into predicate logic in the way outlined in 5.3 would lead to a signature with infinitely many predicate symbols. This supports the intuition that dynamic logic is in fact more powerful than predicate logic, and indeed it can be shown that the set of valid sentences in dynamic logic is not recursively enumerable.

As a consequence, there is no calculus for dynamic logic which is at the same time correct and complete. Nevertheless, correct and "almost-complete" calculi can be designed, and such approaches are of vital importance in modern program verification.

# References

[Schmitt02] Schmitt, P. H.: Nichtklassische Logiken,
Chapter 3: Modale Logiken,
University of Karlsruhe,
`http://i12www.ira.uka.de/studium.htm`

[GaGuen84] Gabbay, D. and Guenthner, F: Handbook of Philosophical Logic,
Volume II: Extensions of Classical Logic,
D. Reidel Publishing Company, 1984