# Modal Logic

CS 3234: Logic and Formal Systems

Martin Henz and Aquinas Hobor

October 6, 2010

Generated on Wednesday 13 October, 2010, 15:58

## 1 Motivation

The source of meaning of formulas in the previous chapters were models. Once a particular model is chosen, say a valuation for propositional logic, the meaning of every formula can be investigated. The formula is considered to hold or not to hold, depending on the given valuation, which assigns truth values $T$ or $F$ to every propositional atom in the formula.

Sometimes, we would like to consider different valuations to be related to each other. For example, we would like to say that a particular proposition is true now, but false in the next second, and remain false from then onwards. In this case, one valuation characterizes the situation now, and another one the situation in one second. Time provides the relationship between these two situations. We can say that a proposition *will always* hold, if it holds in all future situations.

As another example, we may want to describe scenarios that an agent considers believable, when considering certain facts. Here the agent's belief structure defines the relationship between two scenarios. An agent can then be said to *believe* a proposition, given a scenario of facts, if the proposition holds in all scenarios that he considers believable.

Thirdly, we may be interested in necessity and possibility. Here, a scenario is considered possible based on a given situation, if the scenario is consistent with some underlying framework of reality, such as the laws of physics, or philosophical concepts. A proposition would then be called *necessarily* true, if it holds in all possible scenarios.

Ways of reasoning about situations or scenarios are called *modalities*. In the modality of time, we will be able to express that a proposition always holds; in the modality of belief, we will be able to say that an agent believes a proposition; and in the modality of necessity, we will be able to say that a proposition is necessarily true. We shall introduce modal logic to capture such modalities, by first extending the syntax of propositional logic with *modal operators*, and then formally define their semantics, based on situations (or scenarios).

## 2 Syntax

For simplicity, we limit our investigation to a propositional modal logic in this lecture; a study of modal predicate logic is beyond the scope of this module. We add two modal operators to the syntax of propositional logic, namely $\Box$ and $\Diamond$.

**Definition 1.** *For a given set $A$ of propositional atoms, the set of* well-formed formulas in propositional modal logic *is the least set $F$ that fulfills the following rules:*

- *The constant symbols $\bot$ and $\top$ are in $F$.*
- *Every element of $A$ is in $F$.*
- *If $\phi$ is in $F$, then $(\neg\phi)$ is also in $F$.*
- *If $\phi$ and $\psi$ are in $F$, then $(\phi \wedge \psi)$ is also in $F$.*
- *If $\phi$ and $\psi$ are in $F$, then $(\phi \vee \psi)$ is also in $F$.*
- *If $\phi$ and $\psi$ are in $F$, then $(\phi \rightarrow \psi)$ is also in $F$.*
- *If $\phi$ is in $F$, then $(\Box\phi)$ is also in $F$.*
- *If $\phi$ is in $F$, then $(\Diamond\phi)$ is also in $F$.*

Depending on the modality that we are interested in, we may pronounce $\Box\phi$ as:

- $\phi$ holds now and always in the future (modality of time), or
- an agent $A$ believes that $\phi$ holds (modality of belief), or
- $\phi$ necessarily holds (modality of necessity).

Recall that in propositional logic, not all operators are strictly required. We could for example consider $\phi \vee \psi$ as an abbreviation of $\neg(\phi \wedge \phi)$. Thus a propositional logic with all operators except the $\vee$ operator is not less expressive than a propositional logic with all operators. We include $\vee$ for the convenience of succinctly expressing a disjunctive relationship between two formulas.

A similar situation holds for modal logic. We may consider $\Diamond\phi$ to be an abbreviation for $\neg(\Box(\neg\phi))$. For the sake of convenience, we include $\Diamond$ as an operator, although it can be expressed in terms of $\Box$.

Once we have decided to treat $\Diamond\phi$ as an abbreviation for $\neg(\Box(\neg\phi))$, we can work out how to pronounce $\Diamond\phi$ in our example modalities.

- If $\Box\phi$ means that $\phi$ holds now and forever, then $\Diamond\phi \equiv \neg(\Box(\neg\phi))$ means that it is not the case that $\neg\phi$ holds now and forever. That means that there is or will be *some time* at which $\phi$ holds.

- If $\Box\phi$ means that an agent $A$ believes $\phi$, then $\Diamond\phi \equiv \neg(\Box(\neg\phi))$ means that it is not the case that the agent believes $\neg\phi$. In English, we may say that $A$ considers $\phi$ *plausible*.
- If $\Box\phi$ means necessarily $\phi$, then $\Diamond\phi \equiv \neg(\Box(\neg\phi))$ means that it not necessarily the case that $\phi$ does not hold. In other words, $\phi$ may *possibly* hold.

# 3 Semantics

We have seen that the goal of modal logic is to be able to reason about scenarios. In different application domains, these scenarios have different relationships with each other. For example, in the modality of time, a scenario $x$ can be seen to be related to a scenario $y$, written $R(x, y)$, if $y$ occurs at the same time as or a later time than $x$. Since we operate in a propositional setting, each scenario is characterized by a valuation, which is an assignment of propositional variables to $T$ or $F$. In modal logic, we call such scenarios *worlds*.

**Definition 2.** *A* model $\mathcal{M}$ *over a particular set of propositional atoms A is:*

- *a set of worlds* $W$*,*
- *a binary relation* $R \subseteq W \times W$*, called the* accessibility relation*, and*
- *a mapping called* labeling function $L : W \to A \to \{T, F\}$*.*

The following is a lightweight encoding of modal logic using Coq. We start with representing worlds and the accessibility relation.

```
Parameter world : Type.

Parameter R : world -> world -> Prop.
```

Note that we do not specify how a "world" actually looks. A world is entirely described by the operations in which it occurs, namely modal logic propositions.

```
Definition Proposition : Type := world -> Prop.
```

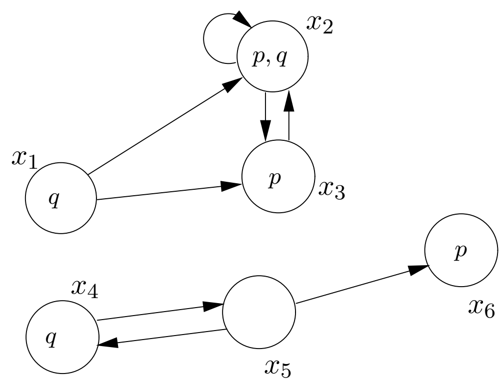
A proposition is a predicate that says whether it holds or does not hold in the world to which it is applied.

Note that for any world $x \in W$, the application $L(w)$ results in a function from $A$ to $\{T, F\}$, in other words a valuation. Models of modal logic that are based on possible worlds are often called *Kripke models* in honour of the logician Saul Kripke, who was instrumental in laying the foundation of modal logic.

**Example 1.** *Consider the following Kripke model, where $L$ lists the propositional atoms that evaluate to $T$ in the respective world.*

$$
\begin{aligned}
W &= \{x_1, x_2, x_3, x_4, x_5, x_6\} \\
R &= \{(x_1, x_2), (x_1, x_3), (x_2, x_2), (x_2, x_3), (x_3, x_2), (x_4, x_5), (x_5, x_4), (x_5, x_6)\} \\
L &= \{(x_1, \{q\}), (x_2, \{p, q\}), (x_3, \{p\}), (x_4, \{q\}), (x_5, \{\}), (x_6, \{p\})\}
\end{aligned}
$$

*We can depict the model graphically as follows.*




Formulas in modal logic do not simply evaluate to $T$ or $F$, but do so relative to a particular world $x$. The following relation $\Vdash$ formalizes this concept.

**Definition 3.** *Let $\mathcal{M} = (W, R, L)$, $x \in W$, and $\phi$ a formula in basic modal logic. For any $x \in W$, we define $x \Vdash \phi$ via structural induction:*

- $x \Vdash \top$ *always holds,*
- $x \Vdash \bot$ *never holds; we write* $x \nVdash \bot$
- $x \Vdash p$ *iff* $p \in L(x)$
- $x \Vdash \neg\phi$ *iff* $x \nVdash \phi$
- $x \Vdash \phi \wedge \psi$ *iff* $x \Vdash \phi$ *and* $x \Vdash \psi$
- $x \Vdash \phi \vee \psi$ *iff* $x \Vdash \phi$ *or* $x \Vdash \psi$
- $x \Vdash \phi \rightarrow \psi$ *iff* $x \Vdash \phi$ *implies that* $x \Vdash \psi$
- $x \Vdash \Box\phi$ *iff for each* $y \in W$ *with* $R(x, y)$, *we have* $y \Vdash \phi$
- $x \Vdash \Diamond\phi$ *iff there is a* $y \in W$ *such that* $R(x, y)$ *and* $y \Vdash \phi$.

Here the definition of $x \Vdash \Diamond\phi$ follows from our definition of $\Diamond$. If $\Diamond\phi$ is an abbreviation for $\neg\Box\neg\phi$, then $x \Vdash \Diamond\phi$ is defined by $x \Vdash \neg\Box\neg\phi$, which means that $x \Vdash \Box\neg\phi$ does not hold. Thus, $\neg\phi$ does not hold for all $y \in W$ with $R(x, y)$. This means that there is at least one $y \in W$ such that $R(x, y)$ and $y \Vdash \phi$.

```
Definition holds_in (w : world)  (phi : Proposition) : Prop :=
  phi w.

Notation "w ||- phi" := (holds_in w phi) (at level 30).
```

To express $w \Vdash \phi$, we apply the predicate `holds_in`, supported by the infix notation `w ||- phi`. Thus, we can use the tactic `unfold holds_in` to apply `phi` to `w` to see if `phi` holds in `w`.

Now we can define the basic (propositional) operators.

```
Definition Top : Proposition :=
  fun w => True.
```

This means `Top` is true in all worlds. Similarly, `Bot` should be false in all worlds.

```
Definition Bot : Proposition :=
  fun w => False.
```

We proceed to define the syntax of propositional modal logic by explaining what propositions result from negation, conjunction, disjunction and implication of other propositions.

```
Definition Neg (phi : Proposition) : Proposition :=
  fun w => ~ (w ||- phi).
Notation "! phi" := (Neg phi) (at level 16).

Definition And (phi psi : Proposition) : Proposition :=
  fun w => (w ||- phi) /\ (w ||- psi).
Notation "phi && psi" := (And phi psi).

Definition Or (phi psi : Proposition) : Proposition :=
  fun w => (w ||- phi) \/ (w ||- psi).
Notation "phi || psi" := (Or phi psi).

Definition Impl (phi psi : Proposition) : Proposition :=
  fun w => (w ||- phi) -> (w ||- psi).
Notation "phi --> psi" := (Impl phi psi) (at level 20, right associativity).
```

Finally, we define the modal operators $\Box$ and $\Diamond$ by using `forall` and `exists` in Coq.

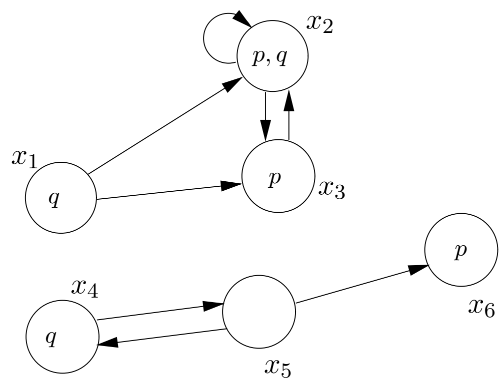
```
Definition Box (phi : Proposition) : Proposition :=
  fun w => forall w', R w w' -> (w' ||- phi).

Notation "[] phi" := (Box phi) (at level 15).

Definition Diamond (phi : Proposition) : Proposition :=
  fun w => exists w', R w w' /\ (w' ||- phi).

Notation "<> phi" := (Diamond phi) (at level 15).
```

**Example 2.** *Consider the same Kripke model as presented in the previous example.*




- $x_1 \Vdash q$
- $x_1 \Vdash \Diamond q$, $x_1 \nVdash \Box q$
- $x_5 \nVdash \Box p$, $x_5 \nVdash \Box q$, $x_5 \nVdash \Box p \vee \Box q$, $x_5 \Vdash \Box(p \vee q)$
- $x_6 \Vdash \Box\phi$ *holds for all* $\phi$, *but* $x_6 \nVdash \Diamond\phi$

*By stating that* `p` *and* `q` *are* `Propositions`*, we are assuming that they either hold in a world, or don't. In other words they behave like propositional atoms.*

```
Section Example.

Parameter p q : Proposition.
```

*After introducing the six worlds*

```
Parameter x1 x2 x3 x4 x5 x6: world.
```

*we proceed to describe $R$ and $L$.*

```
Hypothesis Kripke_example:
                      R x1 x2 /\ R x1 x3 /\ R x2 x3 /\
                      R x3 x2 /\ R x2 x2 /\ R x4 x5 /\
                      R x5 x4 /\ R x5 x6 /\
                      (x1 ||- ! p) /\ (x1 ||- q) /\
                      (x2 ||- p)   /\ (x2 ||- q) /\
                      (x3 ||- p)   /\ (x3 ||- ! q) /\
                      (x4 ||- ! p) /\ (x4 ||- q) /\
                      (x5 ||- ! p) /\ (x5 ||- ! q) /\
                      (x6 ||- p)   /\ (x6 ||- ! q)
                      .
```

*We can show* $x1 \Vdash \Diamond q$ *through* $\exists i$.

```
Lemma Kripke_example_1: x1 ||- <> q.
Proof.
unfold holds_in, Diamond.
exists x2.
tauto.
Qed.




End Example.
```

We said $x_6 \Vdash \Box\phi$ holds for all $\phi$, but $x_6 \nVdash \Diamond\phi$ Greek letters denote formulas, and are not propositional atoms. Terms where Greek letters appear instead of propositional atoms are called *formula schemes*.

**Exercise 1.** *For each of the following formulas, give an example for a Kripke model, in which the formula holds in all worlds of the model.*

- $\Diamond p \wedge \neg\Box p$
- $\Diamond\top$
- $\Box\bot$
- $\Diamond p \wedge \Diamond q \to \Diamond(p \wedge q)$

**Definition 4.** *A set of formulas* $\Gamma$ *entails a formula* $\psi$ *of basic modal logic if, in any world* $x$ *of any model* $\mathcal{M} = (W, R, L)$*, whe have* $x \Vdash \psi$ *whenever* $x \Vdash \phi$ *for all* $\phi \in \Gamma$*. We say* $\Gamma$ *entails* $\psi$ *and write* $\Gamma \models \psi$*.*

We write $\phi \equiv \psi$ if $\phi \models \psi$ and $\psi \models \phi$.
The following list states a few well-known equivalences.

- De Morgan rules: $\neg\Box\phi \equiv \Diamond\neg\phi$, $\neg\Diamond\phi \equiv \Box\neg\phi$.

- Distributivity of $\Box$ over $\wedge$:

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

- Distributivity of $\Diamond$ over $\vee$:

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

- $\Box\top \equiv \top$, $\Diamond\bot \equiv \bot$

**Exercise 2.** *Prove the distributivity of $\Diamond$ over $\vee$ using the definition of $\models$.*

**Definition 5.** *A formula $\phi$ is valid if it is true in every world of every model, i.e. iff $\models \phi$ holds.*

In order to express validity in Coq, we use the definition of $\models$ directly.

```
Definition valid (phi: Proposition) : Prop :=
  forall w, w ||- phi.
Notation "|= phi" := (valid phi) (at level 30).
```

Examples of valid formulas:

- All valid formulas of propositional logic
- $\Diamond\neg\phi \rightarrow \neg\Box\phi$
- $\Box(\phi \wedge \psi) \rightarrow \Box\phi \wedge \Box\psi$
- $\Diamond(\phi \vee \psi) \rightarrow \Diamond\phi \vee \Diamond\psi$
- Formula $K$: $\Box(\phi \rightarrow \psi) \wedge \Box\phi \rightarrow \Box\psi$.

We prove the four formulas above in Coq.

```
Lemma DiamondNotNotBox: forall phi,
  |= <> (! phi) --> ! ([] phi).
Proof.
  intros.
  do 3 intro.
  destruct H.
  unfold holds_in,Box in H0.
  spec H0 x.
  destruct H.
  apply H0 in H.
  contradiction H.
Qed.
```

```
Lemma boxConj: forall phi psi,
|= ([] (phi && psi) --> (([] phi) && ([] psi))).

Proof.
  intros.
  intro.
  intro.
  unfold holds_in, And.
  split.
  intro.
  intro.
  spec H w'.
  apply H in H0.
  destruct H0.
  trivial.
  intro.
  intro.
  spec H w'.
  apply H in H0.
  destruct H0.
  trivial.
Qed.
```

```
Lemma diamondDisj: forall phi psi,
|= (<> (phi || psi) --> ((<> phi) || (<> psi))).
Proof.
intros.
intro.
intro.
destruct H.
destruct H.
unfold holds_in, Or.
unfold holds_in, Or in H0.
destruct H0.
left.
exists x.
auto.
right.
exists x.
auto.
Qed.
```

```
Lemma K : forall phi psi,
  |= ([] (phi --> psi)) --> [] phi --> [] psi.
Proof.
  intros.
  do 5 intro.
  spec H w' H1.
  apply H.
  apply H0.
  trivial.
Qed.
```

The last formula is named K after Saul Kripke. It plays an important role as the modal logic version of modus ponens. in modal logic.

# 4 Logic Engineering

We have seen that in a particular context $\Box\phi$ could mean:

- It will always be true that $\phi$
- Agent $Q$ believes that $\phi$
- It is necessarily true that $\phi$

Other modalities allow us to express:

- It ought to be that $\phi$
- Agent $Q$ knows that $\phi$
- After any execution of program $P$, $\phi$ holds.

Since $\Diamond\phi \equiv \neg\Box\neg\phi$, we can infer the meaning of $\Diamond$ in each context.

| $\Box\phi$ | $\Diamond\phi$ |
|---|---|
| It is necessarily true that $\phi$ | It is possibly true that $\phi$ |
| It will always be true that $\phi$ | Sometime in the future $\phi$ |
| It ought to be that $\phi$ | It is permitted to be that $\phi$ |
| Agent $Q$ believes that $\phi$ | $\phi$ is consistent with $Q$'s beliefs |
| Agent $Q$ knows that $\phi$ | For all $Q$ knows, $\phi$ |
| After any run of $P$, $\phi$ holds. | After some run of P, $\phi$ holds |

## 4.1 Valid Formulas wrt Modalities

In each modality, we can investigate whether particular formulas should be considered valid. A close examination of the modalities allows us to build the following table.

| $\Box\phi$ | $\Box\phi \to \phi$ | $\Box\phi \to \Box\Box\phi$ | $\Diamond\phi \to \Box\Diamond\phi$ | $\Diamond\top$ | $\Box\phi \to \Diamond\phi$ | $\Box\phi \vee \Box\neg\phi$ | $\Box(\phi \to \psi) \to \Box\phi \to \Box\psi$ | $\Diamond\phi \wedge \Diamond\psi \to \Diamond(\phi \wedge \psi)$ |
|---|---|---|---|---|---|---|---|---|
| It will always be that $\phi$ | × | √ | × | × | × | × | √ | × |
| Agent Q believes that $\phi$ | × | √ | √ | √ | √ | × | √ | × |
| It is necessary that $\phi$ | √ | √ | √ | √ | √ | × | √ | × |
| It ought to be that $\phi$ | × | × | × | √ | √ | × | √ | × |
| Agent Q knows that $\phi$ | √ | √ | √ | √ | √ | × | √ | × |
| After running P, $\phi$ | × | × | × | × | × | × | √ | × |

Thus by reasoning from known properties of a modality, we can infer the validity of additional formulas in the modality.

## 4.2 Properties of $R$

We can also infer a "meaning" for the accessibility relation $R$ by investigating the respective modalities. For example, in the modality of time, $R(x, y)$ expresses that $y$ is a future (or present) worlds of $x$.

| $\Box\phi$ | $R(x, y)$ |
|---|---|
| It will always be true that $\phi$ | $y$ is a future world of $x$ |
| Agent Q believes that $\phi$ | $y$ could be the actual world according to Q's beliefs at $x$ |
| It is necessarily true that $\phi$ | $y$ is possible world according to info at $x$ |
| It ought to be that $\phi$ | $y$ is an acceptable world according to the information at $x$ |
| Agent Q knows that $\phi$ | $y$ could be the actual world according to Q's knowledge at $x$ |
| After any execution of P, $\phi$ holds | $y$ is a possible resulting state after execution of P at $x$ |

Recall from the study of discrete mathematics that we can classify binary relations according to their properties.

- reflexive: for every $x \in W$, we have $R(x, x)$.
- symmetric: for every $x, y \in W$, we have $R(x, y)$ implies $R(y, x)$.
- serial: for every $x$ there is a $y$ such that $R(x, y)$.
- transitive: for every $x, y, z \in W$, we have $R(x, y)$ and $R(y, z)$ imply $R(x, z)$.
- Euclidean: for every $x, y, z \in W$ with $R(x, y)$ and $R(x, z)$, we have $R(y, z)$.
- functional: for each $x$ there is a unique $y$ such that $R(x, y)$.

- linear: for every $x, y, z \in W$ with $R(x, y)$ and $R(x, z)$, we have $R(y, z)$ or $y = z$ or $R(z, y)$.
- total: for every $x, y \in W$, we have $R(x, y)$ and $R(y, x)$.
- equivalence: reflexive, symmetric and transitive.

Since we have connnected modalities to the validity of certain formulas on the one hand, and to relations on the other hand, it will come as no surprise that the validity of formulas corresponds to properties of the accessibility relation. This relationship is the basis of *correspondence theory*.

## 4.3 Correspondence Theory

We would like to establish that some formulas hold whenever $R$ has a particular property. In order to investigate this relationship, it will be useful to be able to ignore $L$, and only consider the $(W, R)$ part of a model, which we will call a *frame*. We shall then establish formula schemes based on properties of frames.

**Theorem 1.** *Let $\mathcal{F} = (W, R)$ be a frame. The following statements are equivalent:*

1. *$R$ is reflexive;*
2. *$\mathcal{F}$ satisfies $\Box\phi \to \phi$;*
3. *$\mathcal{F}$ satisfies $\Box p \to p$ for any atom $p$*

*Proof.*

**1 $\Rightarrow$ 2:** Let $R$ be reflexive. Let $L$ be any labeling function; $\mathcal{M} = (W, R, L)$. We need to show for any $x$: $x \Vdash \Box\phi \to \phi$.

Suppose $x \Vdash \Box\phi$. Since $R$ is reflexive, we have $x \Vdash \phi$. Using the semantics of $\to$ we can conclude: $x \Vdash \Box\phi \to \phi$

**2 $\Rightarrow$ 3:** Just set $\phi$ to be $p$

**3 $\Rightarrow$ 1:** Suppose the frame satisfies $\Box p \to p$. Take any world $x$ from $W$.

Choose a labeling function $L$ such that $p \notin L(x)$, but $p \in L(y)$ for all $y$ with $y \neq x$.

Proof by contradiction: Assume $(x, x) \notin R$. Then we would have $x \Vdash \Box p$, but not $x \Vdash p$. Contradiction!

$\Box$

In Coq, the proof is simpler than on paper.

```
Lemma Theorem1:
   (forall w, R w w) <-> forall phi, |= [] phi --> phi.
Proof.
split.
repeat intro.
auto.
repeat intro.
spec H (fun w' => R w w').
apply H.
intro.
intro.
unfold holds_in.
trivial.
Qed.
```

The reverse direction is remarkable, since the proposition, with which the hypothesis `H` is specialized, holds in a world `w'` if and only if `R w w'` holds.

**Theorem 2.** *The following statements are equivalent:*

- *$R$ is transitive;*
- *$\mathcal{F}$ satisfies $\Box\phi \to \Box\Box\phi$;*
- *$\mathcal{F}$ satisfies $\Box p \to \Box\Box p$ for any atom $p$*

**Exercise 3.** *Provide a paper proof for Theorem 2.*

Again, the Coq proof is simple, using a similar trick as for Theorem 1.

```
Lemma Theorem2:
(forall w1 w2 w3, R w1 w2 -> R w2 w3 -> R w1 w3)
<->
(forall phi, |= [] phi --> [] [] phi).
Proof.
split.
repeat intro.
apply H0.
apply H with w'.
apply H1.
apply H2.
intros.
spec H (fun w' => R w1 w') w1.
unfold Box,Impl,holds_in in H.
apply H with w2; auto.
Qed.
```

**Exercise 4.** *Prove (on paper) that $\Box\bot$ corresponds to $R = \emptyset$.*

**Exercise 5.** *Conduct the proof using Coq.*

**Exercise 6.** *Find a property of $R$ that corresponds to the formula $\Diamond\top$. Prove the correspondence on paper and using Coq.*

The following table summarizes important correspondences. Note that names on the left are traditional names of the formulas used in modal logic literature.

| name | formula scheme | property of $R$ |
|---|---|---|
| T | $\Box\phi \to \phi$ | reflexive |
| B | $\phi \to \Box\Diamond\phi$ | symmetric |
| D | $\Box\phi \to \Diamond\phi$ | serial |
| 4 | $\Box\phi \to \Box\Box\phi$ | transitive |
| 5 | $\Diamond\phi \to \Box\Diamond\phi$ | Euclidean |
| | $(\Box\phi \to \Diamond\phi) \wedge (\Diamond\phi \to \Box\phi)$ | functional |
| | $\Box(\phi \wedge \Box\phi \to \psi) \vee \Box(\psi \wedge \Box\psi \to \phi)$ | linear |