Modal Logic:
Relational Semantics and Decidability[*]

Yakov Shklarov

March 2020

Recent decades have seen an explosion of interest in the use of modal logic for formal system verification. Although modal logic is less expressive than first-order logic, it enjoys a key property that makes it well-suited for this task: Unlike first-order logic, modal logic is *decidable*. Roughly speaking, this means that there exist algorithms for determining whether a given modal sentence is true in a given class of models. The aim of this report is to briefly explore the syntax and semantics of modal logic and its relationship to first-order logic, and then to present a proof of the decidability of the validity and satisfiability problems.

We'll limit ourselves to the so-called *basic modal language*. For real-world applications, this is usually insufficient—it's often necessary to impose various systems of axioms, which allow for more powerful proof systems at the cost of restricting the class of models. These axiom systems have standardized names such as "K", "K4", "S5", and "KD4". For more on this topic, see the *Handbook of Modal Logic*, Chapter 2 [1]. Furthermore, one can extend the logic with various non-modal logical connectives—for example, as in temporal logics (see Chapters 11 and 12 of the *Handbook*), which are frequently used in industry to model hardware and software systems.

This report is essentially a more focused version of Chapter 1 of the *Handbook of Modal Logic* [1], with a few proofs explained in more detail. The *Handbook* is a wonderful introduction to many diverse aspects of the field. For a more elementary treatment, see Blackburn, de Rijke, and Venema's *Modal Logic* [2]. For more on formal system verification, see the references on model checking [3] [4] [5] [6].

Contents

[*] Submitted to Christopher Eagle in partial fulfillment of the requirements for the undergraduate course "Math 493: Model Theory" at the University of Victoria.

1 The Language and Its Relational Models

There are several different semantics for modal logic. We will work with the *relational* (or *Kripke*) semantics, which is the most widely used.

In classical model theory, the signature of a language is made up of symbols that get interpreted as constants, functions, and relations in an algebraic structure. In modal logic, the model is a very specific kind of algebraic structure: a graph with labels on the vertices and edges. The labels of the vertices and edges comprise the signature of the language. Every sentence takes on a truth value at each vertex of the graph (not in the model as a whole).

Definition. *A* ***signature*** *is a pair $(MOD, PROP)$ where MOD is a nonempty set and PROP is a set. The elements of MOD are called* ***modality symbols****. The elements of PROP are called* ***propositional symbols****.*

We will often implicitly assume that the signature has been fixed.

We have to describe three things: the syntax of the language (what the formulas are), its relational semantics (what the models are), and the truth schema (what it means for a formula to be true in a model).

Definition. *The* ***(basic) modal language*** *with signature $(MOD, PROP)$ is the set of expressions (called* ***formulas****) defined recursively as follows:*

- *For every propositional symbol $p \in PROP$, p is a formula.*
- *The expressions $\bot$ and $\top$ are formulas.*
- *Whenever φ and ψ are formulas, the following are also formulas: $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \to \psi$, $\varphi \leftrightarrow \psi$.*
- *Whenever φ is a formula, for every modality symbol $m \in MOD$, the expressions $\Box_m\varphi$ and $\Diamond_m\varphi$ are formulas.*

If MOD has only one element, then the subscript will be omitted: $\Box\varphi$, $\Diamond\varphi$.

Note that there's no distinction between formulas and sentences, because there's no notion of a variable (free or otherwise). The box and diamond connectives let us do a very limited kind of quantification. This idea will be formalized in section 3.

Definition. *A (****relational*** *or* ***Kripke****)* ***model*** *over a signature $(MOD, PROP)$ is a triple $\mathcal{W} = (W, \{R^m\}_{m \in MOD}, V)$ where:*

- *The* ***universe*** *W is a nonempty set whose elements are called* ***points*** *(or, in certain contexts,* ***worlds*** *or* ***states****).*
- *For each $m \in MOD$, R^m is a binary relation on W. (That is, (W, R^m) is a digraph, possibly with loops.)*
- *The* ***valuation*** *V is a map from PROP to $\mathcal{P}(W)$. (Each propositional symbol gets attached to zero or more points.)*

A ***pointed model*** *is an ordered pair $(\mathcal{M}, w)$ where $\mathcal{M}$ is a model and w is a point of $\mathcal{M}$. It's common to omit the brackets and write simply $\mathcal{M}, w$.*

A model (or pointed model) is called ***finite*** *if its universe is finite.*

Definition. *Let $\mathcal{W} = (W, \{R^m\}_{m \in MOD}, V)$ be a signature, $\mathcal{M}$ a model, and φ a formula. We say that φ is* ***true*** *in $\mathcal{M}$ at the point $w \in W$, or that $\mathcal{M}$* ***satisfies*** *φ at w, and write*

$$\mathcal{M}, w \models \varphi,$$

if the following recursive truth schema is satisfied.

- $\mathcal{M}, w \models p$ *for $p \in PROP$ if and only if $w \in V(p)$ (that is, the point w has label p).*
- $\mathcal{M}, w \not\models \bot$.
- $\mathcal{M}, w \models \top$.
- $\mathcal{M}, w \models \neg\varphi$ *if and only if* $\mathcal{M}, w \not\models \varphi$.

- $\mathcal{M}, w \models \varphi \wedge \psi$ *if and only if both* $\mathcal{M}, w \models \varphi$ *and* $\mathcal{M}, w \models \psi$.

- $\mathcal{M}, w \models \varphi \vee \psi$ *if and only if either* $\mathcal{M}, w \models \varphi$ *or* $\mathcal{M}, w \models \psi$.

- $\mathcal{M}, w \models \varphi \to \psi$ *if and only if* $\mathcal{M}, w \models \neg\varphi \vee \psi$.

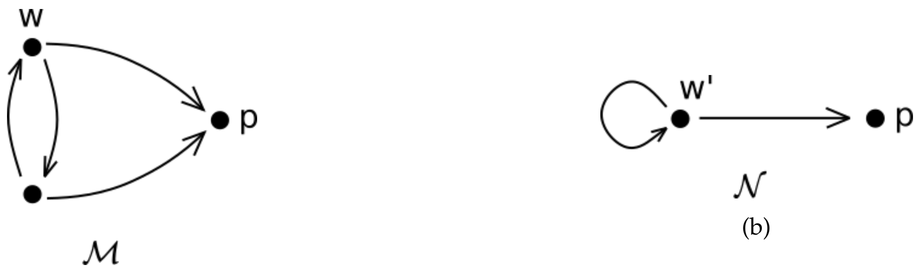
- $\mathcal{M}, w \models \varphi \leftrightarrow \psi$ *if and only if* $\mathcal{M}, w \models (\varphi \wedge \psi) \vee (\neg\varphi \wedge \neg\psi)$.

- $\mathcal{M}, w \models \Box_{\mathfrak{m}} \varphi$ *if and only if for every* $v \in W$ *such that* $R^{\mathfrak{m}} wv$ *we have* $\mathcal{M}, v \models \varphi$. *(That is,* φ *is true at every* $\mathfrak{m}$*-successor.)*

- $\mathcal{M}, w \models \Diamond_{\mathfrak{m}} \varphi$ *if and only if there exists* $v \in W$ *such that* $R^{\mathfrak{m}} wv$ *and* $\mathcal{M}, v \models \varphi$. *(That is,* φ *is true at some* $\mathfrak{m}$*-successor.)*

From now on, unless otherwise specified, the logic will be assumed to be **unimodal** in the sense that there is only one modality symbol (one kind of edge; one kind of box/diamond); the opposite case is called **multimodal**. It should be mentioned that it's also possible to allow boxes and diamonds to have arity greater than 1: such connectives are called **polyadic modalities**.

We will also assume that there are at most countably many propositional symbols.



(a)

(b)

Figure 1: Two pointed models

Here are two examples. The pointed model $(\mathcal{M}, w)$ is depicted in fig. 1a; the pointed model $(\mathcal{N}, w')$ is depicted in fig. 1b. In each case, a single point is labeled with the propositional symbol p. We have

$$\begin{aligned}
&\mathcal{M}, w \models \Diamond p, \\
&\mathcal{M}, w \models \Diamond\Diamond p, \\
&\mathcal{M}, w \not\models \Box p, \\
&\mathcal{M}, w \models \Diamond(p \wedge \Box\bot).
\end{aligned}$$

The formula $\Box\bot$ is true at a point if and only if there no outgoing edges; otherwise, $\Diamond\top$ is true.

The four formulas above are also all true in $(\mathcal{N}, w')$. In fact, we'll prove in the next section that $(\mathcal{M}, w)$ and $(\mathcal{N}, w')$ satisfy precisely the same modal formulas. So the situation is very different from first-order logic, where the language can always distinguish between two finite non-isomorphic structures.

Observe that for every pointed model $(\mathcal{M}, w)$ and every formula φ

$$\mathcal{M}, w \models \neg\Box\varphi \quad \Longleftrightarrow \quad \mathcal{M}, w \models \Diamond\neg\varphi.$$

Thus, every formula is equivalent to a formula that uses only the four logical connectives $\bot, \neg, \wedge, \Box$, together with propositional symbols. This observation will simplify proofs that use induction on formulas.

2 Modal Equivalence and Bisimulation

Isomorphism of models is simply isomorphism of labeled graphs.

Definition. *Given two models $\mathcal{M} = (W, \{R^m\}_{m \in MOD}, V)$ and $\mathcal{N} = (W', \{R'^m\}_{m \in MOD}, V')$, an **isomorphism** from $\mathcal{M}$ to $\mathcal{N}$ is a bijection $\eta : W \to W'$ such that $R^m vw \iff R'^m \eta(v)\eta(w)$ for every $m \in MOD$ and every $v, w \in W$, and also $\eta(V(p)) = V'(p)$ for every $p \in PROP$. Given two pointed models $(\mathcal{M}, s)$ and $(\mathcal{N}, t)$, an **isomorphism** is an isomorphism η from $\mathcal{M}$ to $\mathcal{N}$ such that $\eta(s) = t$.*

*Two models (or pointed models) are called **isomorphic** if there exists an isomorphism from one to the other.*

In first-order model theory we can have *elementary equivalence* of structures; in modal logic the analogous concept is *modal equivalence*.

Definition. *Two pointed models $(\mathcal{M}, s)$ and $(\mathcal{N}, t)$ are called **modally equivalent** if, for every modal formula φ,*

$$\mathcal{M}, s \models \varphi \quad \textit{if and only if} \quad \mathcal{N}, t \models \varphi.$$

There is also a semantic notion called *bisimulation* that is closely related to modal equivalence. Bisimulation unifies several older constructions in modal logic: generated submodels, p-morphisms, and disjoint unions [1]. The definition was first formulated by logician Johan van Benthem in 1976 (and also independently discovered a few years later in the theory of automata in computer science [1]). Many proofs rely on bisimulations, as for example in section 4 and section 5.

Definition. *A **bisimulation** between pointed models*

$$(\mathcal{M}, s) = (W, \{R^m\}_{m \in MOD}, V, s), \qquad (\mathcal{N}, t) = (W', \{R'^m\}_{m \in MOD}, V', t)$$

is a binary relation B between W and W' that satisfies

- $(s, t) \in B$,
- *For every $(w, w') \in B$ and every $p \in PROP$, we have $w \in V(p)$ if and only if $w' \in V'(p)$ (**atomic harmony**),*
- *If $(w, w') \in B$ and $R^m wv$, then there exists $v' \in W'$ such that $(v, v') \in B$ and $R'^m w'v'$ (**zig**),*
- *If $(w, w') \in B$ and $R'^m w'v'$, then there exists $v \in W$ such that $(v, v') \in B$ and $R^m wv$ (**zag**).*

The definition is symmetric: If we swap the two models, zig becomes zag and zag becomes zig.

As an example, recall the two models in fig. 1. Let B be the binary relation that relates both of the left-hand points in $\mathcal{M}$ to w', and relates the rightmost point of $\mathcal{M}$ to the rightmost point of $\mathcal{N}$. Then B is a bisimulation.

Modal formulas can be thought of as macros or automata that begin at the selected point of the model and execute step-by-step according to the relation(s) of the model. For this reason, the relations are sometimes called **accessibility relations**: They encode when a point is accessible from another. The definition of bisimulation is designed so that formulas can't tell which model they're in—from the perspective of modal logic, bisimilar models look identical.

Theorem 1. *If there exists a bisimulation between two pointed models, then they are modally equivalent.*

Proof. We'll use induction on formulas. Let φ be a formula and assume that the result holds for every subformula of φ. Let B be a bisimulation between

$$(\mathcal{M}, s) = (W, \{R^m\}_{m \in \mathrm{MOD}}, V, s) \quad \text{and} \quad (\mathcal{N}, t) = (W', \{R'^m\}_{m \in \mathrm{MOD}}, V', t).$$

By the comment at the end of section 1, it suffices to check five cases.

i. If $\varphi = p$ for some $p \in \mathrm{PROP}$, then

$$\mathcal{M}, s \models \varphi \iff s \in V(p) \iff t \in V'(p) \iff \mathcal{N}, t \models \varphi.$$

ii. If $\varphi = \bot$, then $\mathcal{M}, s \not\models \varphi$ and $\mathcal{N}, t \not\models \varphi$.

iii. If $\varphi = \neg\psi$, then

$$\mathcal{M}, s \models \varphi \iff \mathcal{M}, s \not\models \psi \iff \mathcal{N}, t \not\models \psi \iff \mathcal{N}, t \models \varphi.$$

iv. If $\varphi = \psi \wedge \xi$, then

$$\mathcal{M}, s \models \varphi \iff \mathcal{M}, s \models \psi \text{ and } \mathcal{M}, s \models \xi \iff \mathcal{N}, t \models \psi \text{ and } \mathcal{N}, t \models \xi \iff \mathcal{N}, t \models \varphi.$$

v. If $\varphi = \Box_m \psi$, then

$$\mathcal{M}, s \models \varphi \iff \mathcal{M}, w \models \psi \text{ whenever } R^m sw.$$

If $w' \in W'$ and $R'^m tw'$, then by zag (since $(s, t) \in B$) we can take $w \in W$ such that $R^m sw$ and $(w, w') \in B$. If $\mathcal{M}, w \models \psi$ then $\mathcal{N}, w' \models \psi$. So in fact

$$\mathcal{M}, s \models \varphi \iff \mathcal{M}, w \models \psi \text{ whenever } R^m sw \implies \mathcal{N}, w' \models \psi \text{ whenever } R'^m tw' \iff \mathcal{N}, t \models \varphi.$$

And, since the converse of B is also a bisimulation, $\mathcal{N}, t \models \varphi \implies \mathcal{M}, s \models \varphi$.

□

There is also a partial converse.

Definition. *A model $\mathcal{M} = (W, \{R^m\}_{m \in MOD}, V)$ has* ***finite branching width*** *if for every $w \in W$ and every $m \in MOD$, there are only finitely many $v \in W$ such that $R^m wv$.*

Theorem 2. *If two pointed models with finite branching width are modally equivalent, then there exists a bisimulation between them.*

Proof. Let

$$(\mathcal{M}, s) = (W, \{R^m\}_{m \in \text{MOD}}, V, s) \quad \text{and} \quad (\mathcal{N}, t) = (W', \{R'^m\}_{m \in \text{MOD}}, V', t)$$

be modally equivalent pointed models with finite branching width. Define

$$B := \{(w, w') \in W \times W' \mid (\mathcal{M}, w) \text{ and } (\mathcal{N}, w') \text{ are modally equivalent}\}.$$

Then B is a bisimulation between $(\mathcal{M}, s)$ and $(\mathcal{N}, t)$:

1. By assumption, $(s, t) \in B$.
2. Atomic harmony: Assume $(w, w') \in B$. If $p \in$ PROP and $w \in V(p)$, then $\mathcal{M}, w \models p$ and hence $\mathcal{N}, w' \models p$, so $w' \in V'(p)$. Similarly, $w' \in V'(p)$ implies $w \in V(p)$.
3. Zig: Assume $(w, w') \in B$ and $R^m wv$ for some $m \in$ MOD. Suppose for the sake of contradiction that for every $v' \in W'$ with $R'^m w'v'$, the points v and v' do not satisfy the same modal formulas. It follows that for every such v' there exists a formula that is true at v' but not at v (because if ψ is true at v but not v', then $\neg\psi$ is true at v' but not v). Pick a single such formula for every R'^m-successor of w', and enumerate these formulas as $\varphi_1, \dots, \varphi_n$. We have

   $$\mathcal{N}, w' \models \Box_m(\varphi_1 \vee \cdots \vee \varphi_n)$$

   and therefore

   $$\mathcal{M}, w \models \Box_m(\varphi_1 \vee \cdots \vee \varphi_n).$$

   It follows that

   $$\mathcal{M}, v \models \varphi_1 \vee \cdots \vee \varphi_n.$$

   But this contradicts the assumption that $\varphi_1, \dots, \varphi_n$ are all false at v.
4. Zag: The definition of B is symmetric, so the argument for zig also proves zag.

□

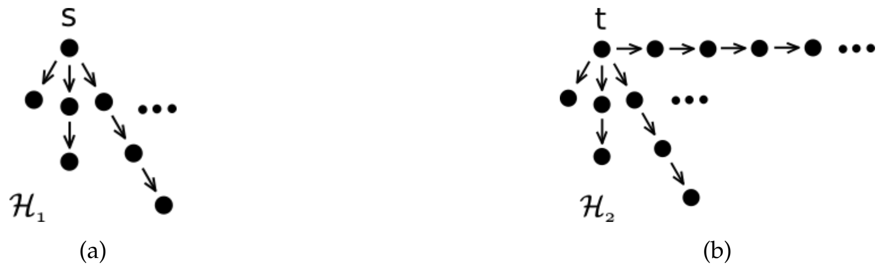


Figure 2: Two hedgehogs

The assumption of finite branching width cannot be removed. For example, fig. 2 depicts two "hedgehog" models: Let PROP $= \varnothing$ and MOD $= \{m\}$, and let $(\mathcal{H}_1, s)$ be the model that branches from s into countably many chains: a chain of length 1, a chain of length 2, ..., a chain of length n for every $n \in \mathbb{N}_{\geqslant 1}$. Let $(\mathcal{H}_2, t)$ be the hedgehog with an additional countably infinite chain branching out from t. The same modal formulas are true at s and t because every formula has finite length and so cannot distinguish the presence of the infinite chain: anything that is true by traveling along the infinite chain will also be true by traveling along a sufficiently long finite chain. But the two hedgehogs are not bisimilar, because the first node along the infinite chain in $\mathcal{H}_2$ has no zag partner in $\mathcal{H}_1$—some formula of the form $\Box\Box\cdots\Box\bot$ will distinguish this first node from any successor of s.

However, we can always get a bisimulation if we're willing to move up to an elementary extension (in the first-order sense). To make this statement precise, we first have to describe how to translate between the modal language and the first-order language: this will be the subject of section 3.

But before leaving this section, we'll define a certain construction that is necessary for one of the proofs in section 4. The idea is to convert an arbitrary model to a (possibly infinite) rooted tree that is bisimilar to the original model. This *tree unraveling* is valuable as a proof device because trees are often easier to work with than arbitrary graphs.

Definition. *Let* $(\mathcal{M}, s) = (W, \{R^m\}_{m \in MOD}, V, s)$ *be a pointed model. The* ***tree unraveling*** *of* $(\mathcal{M}, s)$ *is the model* $(\mathcal{N}, t) = (W', \{R'^m\}_{m \in MOD}, V', t)$ *whose universe* W' *consists of all finite directed walks (i.e., finite sequences of alternating vertices and edges that connect them) in* $(\mathcal{M}, s)$ *that begin with* s*, where* $R'^m \alpha\beta$ *for* $\alpha, \beta \in W'$ *if and only if the walk* β *is obtained from* α *by traveling one* m*-edge further. The propositional symbols that are true at* α *are those that are true at the last vertex of* α*. And* $t = (s)$*.*

For example, the tree unraveling of the two-point model with a loop (fig. 1b) is isomorphic to $\mathcal{H}_2$ (fig. 2a) after labeling each quill tip of the hedgehog with p.

Theorem 3. *The tree unraveling of a model is bisimilar to the original model. One bisimulation is the projection of a walk onto its last vertex.*

Proof. There's nothing to prove: Each of the four conditions for bisimulation is captured directly by the definition of tree unraveling. □

3 The Standard Translation to First-Order Logic

Relational models can be viewed as algebraic structures as in first-order model theory. Modal formulas get translated into first-order formulas. To do this, we must describe the signature of the first-order language.

Definition. *Let* $(PROP, MOD)$ *be a signature. The* ***first-order correspondence language*** *is the first-order language with a binary relation symbol* R^m *for each* $m \in MOD$*, a unary relation symbol* p *for each* $p \in PROP$*, and no constant or function symbols.*

A relational model $\mathcal{M} = (W, \{R^m\}_{m \in \mathrm{MOD}}, V)$ can be viewed as an algebraic structure over the first-order correspondence language, where for every $p \in$ PROP the interpretation of p is $V(p)$ and for every $m \in$ MOD the interpretation of m is R^m. We will denote by $\mathcal{M}$ both the relational model and its corresponding algebraic structure.

Definition. *Let φ be a modal formula. The **standard translation** of φ in the free variable x, denoted $ST_x\varphi$, is the first-order formula defined recursively as follows:*

- $ST_x p := p(x)$ *for every* $p \in PROP$
- $ST_x \bot := x \neq x$
- $ST_x \top := x = x$
- $ST_x \neg\psi := \neg ST_x \psi$
- $ST_x(\psi \wedge \xi) := (ST_x\psi) \wedge (ST_x\xi)$
- $ST_x(\psi \vee \xi) := (ST_x\psi) \vee (ST_x\xi)$
- $ST_x(\psi \to \xi) := ST_x\psi \to ST_x\xi$
- $ST_x(\psi \leftrightarrow \xi) := (ST_x\psi \wedge ST_x\xi) \vee (\neg ST_x\psi \wedge \neg ST_x\xi)$
- $ST_x(\Box_m\psi) := \forall y\ (R^m(x,y) \to ST_y\psi)$ *for every* $m \in MOD$
- $ST_x(\Diamond_m\psi) := \exists y\ (R^m(x,y) \wedge ST_y\psi)$ *for every* $m \in MOD$

It's fairly intuitive that the standard translation is compatible with the truth schemas, but we'll sketch a proof anyway.

Theorem 4. *Let $\mathcal{M} = (W, \{R^m\}_{m\in MOD}, V)$ be a relational model over signature $(MOD, PROP)$. Let φ be a modal formula. For every $w \in W$,*

$$\mathcal{M}, w \models \varphi \quad \textit{if and only if} \quad \mathcal{M} \models ST_w\varphi,$$

where the first turnstile is interpreted according to the modal truth schema, and the second turnstile is interpreted according to the first-order (Tarski) truth schema.

Proof. We use induction on modal formulas.

- For $p \in$ PROP,

$$\mathcal{M}, w \models p \iff w \in V(p) \iff \mathcal{M} \models p(w) \iff \mathcal{M} \models \mathrm{ST}_w p.$$

- For $m \in$ MOD,

$$\begin{aligned}\mathcal{M}, w \models \Box_m\psi &\iff \text{for every } v \in W \text{ with } R^m wv,\ \mathcal{M}, v \models \psi\\ &\iff \text{for every } v \in W \text{ with } R^m wv,\ \mathcal{M} \models \mathrm{ST}_w\psi\\ &\iff \mathcal{M} \models \forall v\ (R^m wv \to \mathrm{ST}_w\psi)\\ &\iff \mathcal{M} \models \mathrm{ST}_w(\Box_m\psi).\end{aligned}$$

The other cases are similar. □

The standard translation lets us get many theorems for free. For example, modal logic has the Löwenheim–Skolem property and the Compactness property. We'll prove just the Compactness property.

Theorem 5 (Compactness). *Let F be a set of modal formulas. If, for every finite subset $\Delta \subseteq F$, there exists a pointed model $(\mathcal{M}, s)$ such that*

$$\mathcal{M}, s \models \varphi \quad \textit{for every } \varphi \in \Delta,$$

then there exists a pointed model $(\mathcal{N}, t)$ such that

$$\mathcal{N}, t \models \varphi \quad \textit{for every } \varphi \in F.$$

Proof. Let $F' = \{\mathrm{ST}_x\varphi \mid \varphi \in F\}$, and define Δ' similarly. Expand the first-order correspondence language by adding a single constant symbol s. By assumption, for every finite $\Delta' \subseteq F'$, there exists a structure $\mathcal{M}$ (in the expanded language) such that

$$\mathcal{M} \models \varphi'(s) \quad \text{for every } \varphi' \in \Delta'.$$

By the Compactness property of first-order logic, it follows that there exists a structure $\mathcal{N}$ such that

$$\mathcal{N} \models \varphi'(s) \quad \text{for every } \varphi' \in F'.$$

Let t be the interpretation of s in $\mathcal{N}$. The conclusion follows by theorem 4. □

4 The Lifting Lemmas

We've seen a couple of notions of equivalence for relational models. Now we'll mention something of their relationship to first-order notions of equivalence.

Theorem 6. *(First Lifting Lemma) Two pointed models* $(\mathcal{M}, s)$ *and* $(\mathcal{N}, t)$ *are modally equivalent if and only if there exist elementary extensions (in the first-order sense) to pointed models* $(\mathcal{M}^+, s)$ *and* $(\mathcal{N}^+, t)$ *that are bisimilar.*

Theorem 7. *(Second Lifting Lemma) Two pointed models* $(\mathcal{M}, s)$ *and* $(\mathcal{N}, t)$ *are modally equivalent if and only if there exist models* $(\mathcal{M}^*, s)$ *and* $(\mathcal{N}^*, t)$*, bisimilar to* $(\mathcal{M}, s)$ *and* $(\mathcal{N}, t)$ *respectively, such that* $(\mathcal{M}^*, s)$ *and* $(\mathcal{N}^*, t)$ *are elementarily equivalent (in the first-order sense).*

A picture of the Lifting Lemmas is presented in fig. 3.

$$\begin{tikzcd}
(\mathcal{M}^+, s) \arrow[r, leftrightsquigarrow] & (\mathcal{N}^+, t) \\
(\mathcal{M}, s) \arrow[u, hook, "\preceq"' sloped] \arrow[r, leftrightarrow, "\equiv_{\Box}"] \arrow[d, leftrightsquigarrow] & (\mathcal{N}, t) \arrow[u, hook, "\preceq"' sloped] \arrow[d, leftrightsquigarrow] \\
(\mathcal{M}^*, s) \arrow[r, leftrightarrow, "\equiv"] & (\mathcal{N}^*, t)
\end{tikzcd}$$

Figure 3: The two Lifting Lemmas

Proving the Lifting Lemmas is rather involved. To keep this report to a sane length, the proofs have been omitted.

A proof of the First Lifting Lemma may be found on pages 115-116 of Maarten de Rijke's doctoral thesis *Extending Modal Logic* [7] (the lemma was first proved in 1976 by Johan van Benthem). The proof entails taking ultrapowers of the models and showing that they are ω-saturated.

A proof of the Second Lifting Lemma may be found on pages 229-232 of Andréka, Németi, and van Benthem's *Modal Languages and Bounded Fragments of Predicate Logic* [8] (the same reference has much, much more to say about the interplay between modal logic and first-order logic.) The proof uses tense logic, a modified version of tree unraveling, and an intricate Ehrenfeucht–Fraïssé argument.

5 Filtrations and the Finite Model Property

In first-order logic, there exist finitely axiomatizable satisfiable theories with no finite models: For example, the theory of dense linear orders without endpoints. In modal logic, things are quite different: If a modal formula has a model, then it must have a finite model. This will be a key fact in our proof of decidability in section 6.

We will obtain a finite model by taking a certain quotient called a *filtration*. Filtrations are a common device in modal logic, often used for proving decidability and completeness results.

Definition. *Let* $\mathcal{M} = (W, R, V)$ *be a model (with a single modality). Let* Σ *be a set of modal formulas closed under subformulas (in the sense that if* $\varphi \in \Sigma$ *and if* ψ *is a subformula of* φ*, then* $\psi \in \Sigma$*). Define an equivalence relation on the points of* W*: Let* $w \leftrightsquigarrow_{\Sigma} w'$ *if and only if the same formulas from* Σ *are satisfied at* w *and at* w'*. Let* $[w]$ *denote the equivalence class of* w*, and let* W_{Σ} *be the set of equivalence classes.*

A ***filtration*** *of* $\mathcal{M}$ *through* Σ *is a model* $\mathcal{M}^f = (W_{\Sigma}, R^f, V^f)$ *such that*

- *If* Rwv *then* $R^f[w][v]$*,*
- *If* $R^f[w][v]$ *and* $\Diamond\varphi \in \Sigma$*, then the condition* $\mathcal{M}, v \models \varphi$ *implies* $\mathcal{M}, w \models \Diamond\varphi$*, and*
- *For every* $p \in$ *PROP,* $V^f(p) = \{[w] \mid w \in V(p)\}$*.*

Notice that if Σ is a subformula-closed set of formulas, then a filtration through Σ can have at most $2^{|\Sigma|}$ points, because the equivalence classes are determined by truth-value assignments to elements of Σ. In particular, if Σ is finite then every filtration is finite.

It needs to be shown that a filtration actually exists. There are typically many filtrations on any given model, and there is a lot to say about this topic [1] [2], but for our purposes it's enough to just exhibit one.

Theorem 8. *For every model $\mathcal{M}$ and subformula-closed set of formulas Σ, there exists a filtration of $\mathcal{M}$ through Σ.*

Proof. Define V^f as in the definition of a filtration, and let

$$R^f[w][v] \text{ if and only if } Rw'v' \text{ for some } w' \in [w] \text{ and some } v' \in [v].$$

The condition $Rwv \implies R^f[w][v]$ is clearly satisfied. As for the other condition, assume that $R^f[w][v]$ and $\Diamond\varphi \in \Sigma$, and that $\mathcal{M}, v \models \varphi$. Take $w' \in [w]$ and $v' \in [v]$ such that $Rw'v'$. Since $v \leftrightsquigarrow_\Sigma v'$ and Σ is closed under subformulas, we have $\mathcal{M}, v' \models \varphi$ and therefore $\mathcal{M}, w' \models \Diamond\varphi$. Since $w \leftrightsquigarrow_\Sigma w'$, it follows that $\mathcal{M}, w \models \Diamond\varphi$, as required. □

The definition of a filtration is designed for the following theorem.

Theorem 9 (Filtration Theorem)**.** *Let $\mathcal{M} = (W, R, V)$ be a model, let Σ be a subformula-closed set of formulas none of which contain $\Box$, and let $\mathcal{M}^f = (W_\Sigma, R^f, V^f)$ be a filtration of $\mathcal{M}$ through Σ. For every $w \in W$ and every $\varphi \in \Sigma$, we have*

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}^f, [w] \models \varphi.$$

Proof. We can do structural induction on formulas, because Σ is closed under subformulas. Assume that $\varphi \in \Sigma$.

The base cases $\varphi = \bot$ and $\varphi = p$ are trivial. If $\varphi = \neg\psi$, then

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}, w \not\models \psi \iff \mathcal{M}^f, [w] \not\models \psi \iff \mathcal{M}^f, [w] \models \varphi.$$

If $\varphi = \psi \wedge \xi$, then

$$\mathcal{M}, w \models \varphi \iff \mathcal{M}, w \models \psi \text{ and } \mathcal{M}, w \models \xi \iff \mathcal{M}^f, [w] \models \psi \text{ and } \mathcal{M}^f, [w] \models \xi \iff \mathcal{M}^f, [w] \models \varphi.$$

The cases $\varphi = \top$, $\varphi = \psi \vee \xi$, $\varphi = \psi \to \xi$, and $\varphi = \psi \leftrightarrow \xi$ are similar.

If $\varphi = \Diamond\psi$ and $\mathcal{M}, w \models \varphi$, then there exists $v \in W$ such that Rwv and $\mathcal{M}, v \models \psi$. By the inductive hypothesis, $\mathcal{M}^f, [v] \models \psi$. But $R^f[w][v]$, so $\mathcal{M}^f, [w] \models \Diamond\psi$.

Conversely, if $\varphi = \Diamond\psi$ and $\mathcal{M}^f, [w] \models \varphi$, then $\mathcal{M}^f, [v] \models \psi$ for some $v \in W$ such that $R^f[w][v]$. By the inductive hypothesis, $\mathcal{M}, v \models \psi$, and by the definition of filtration it follows that $\mathcal{M}, w \models \Diamond\psi$. □

Note that the case $\varphi = \Box\psi$ cannot be dealt with in a straightforward manner: although every formula is equivalent to a formula with $\Diamond$s instead of $\Box$s, the formulas with $\Diamond$s may fail to belong to Σ. The references [1] [2] do not correctly deal with this issue. It may be more elegant to impose a different condition, for example, that Σ be closed under syntactically equivalent formulas, or to strengthen the definition of a filtration.

Now, the main theorem of this section.

Theorem 10 (Strong Finite Model Property)**.** *Let φ be a modal formula with no $\Box$s (as we've seen in section 1, every modal formula is equivalent to such a formula). If φ is satisfied at some point in some (possibly infinite) model, then φ is satisfied at some point in some model with at most $2^{s(\varphi)}$ points, where $s(\varphi)$ is the number of subformulas of φ.*

Proof. Assume that $\mathcal{M}, w \models \varphi$, and let Σ be the set of all subformulas of φ. Let $\mathcal{M}^f$ be a filtration of $\mathcal{M}$ through Σ. Then $\mathcal{M}^f, [w] \models \varphi$, and $\mathcal{M}^f$ has at most $2^{s(\varphi)}$ points. □

6 Decidability and Model Checking

Two basic questions can be asked about a formula.

Definition. *Let φ be a modal formula. We say that φ is* ***satisfiable*** *if there's some model $\mathcal{M}$ and some point w in $\mathcal{M}$ such that $\mathcal{M}, w \models \varphi$. We say that φ is* ***valid*** *if $\mathcal{M}, w \models \varphi$ for every model $\mathcal{M}$ and every point w in $\mathcal{M}$.*

There are many variations on the theme: Given a formula φ, we might be interested, for instance, in finding the set of points in a given model where φ is true. Or we might have a set Σ of formulas (i.e., axioms), and ask whether there exists a model such that every element of Σ is true at every point, and φ is true in at least one point (this is called the **local-global** satisfiability task.)

Evidently, satisfiability and validity are dual concepts: A formula is valid if and only if its negation is not satisfiable. So if an algorithm exists for determining whether a formula is satisfiable, then an algorithm also exists for determining whether a formula is valid. In fact, such an algorithm does exist, so we say that modal logic is **decidable**.

Theorem 11. *There exists an algorithm that, given a modal formula φ, decides whether φ is satisfiable.*

Informal proof sketch. Let φ be a formula, and let ψ be an equivalent formula that has no $\Box$s. By the Strong Finite Model Property, if φ is satisfiable then it's satisfiable at some point in some model with at most $2^{s(\psi)}$ points. But there are only finitely many such models, because once the number of points has been fixed there are only finitely many ways to define R^m and $V(p)$ for those elements $m \in \text{MOD}$ and $p \in \text{PROP}$ that appear within φ. After enumerating these models, the truth value of φ can be checked at each point in each model.

To check the value of φ at a particular point in a particular finite model, we can use a bottom-up labeling algorithm: Build up to φ recursively via subformulas. When working with the subformula ψ, iterate over all points and add the label "ψ" to each point where ψ is true. □

Other proofs of decidability exist: For instance, it's possible to identify modal logic with a certain fragment of first-order logic that has the decidability property [8].

In industrial applications, a real-world system is formalized as a relational model, or a class of such systems is formalized as a system of axioms. Then various desirable properties—criteria for performance and safety of the system—are encoded as modal formulas. A computer program then runs the algorithm to make sure that the properties are always satisfied. This is useful in systems where safety is critical, or when the system is too complicated to be adequately tested via other methods. For example, this methodology has been applied to train control systems to ensure that crashes do not occur. It's also routinely used during the design of microprocessors.

The algorithm described above is only of theoretical interest: it's far too inefficient to be useful in practice. There has been significant progress in the development of efficient algorithms for model checking, and now models with an astronomical number of states can be readily checked [9].

References

[1] P. Blackburn, J. van Benthem, and F. Wolter, eds., *Handbook of modal logic*, vol. 3 of *Studies in Logic and Practical Reasoning*. Elsevier B. V., Amsterdam, 2007.

[2] P. Blackburn, M. de Rijke, and Y. Venema, *Modal logic*, vol. 53 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge, 2001.

[3] E. M. Clarke, Jr., O. Grumberg, and D. A. Peled, *Model Checking*. The MIT Press, 1999.

[4] E. M. Clarke, Jr., O. Grumberg, D. Kroening, D. Peled, and H. Veith, *Model checking*. The Cyber-Physical Systems Series, MIT Press, Cambridge, MA, second ed., 2018. With a foreword by Amir Pnueli.

[5] C. Baier and J.-P. Katoen, *Principles of model checking*. MIT Press, Cambridge, MA, 2008. With a foreword by Kim Guldstrand Larsen.

[6] M. Huth and M. Ryan, *Logic in Computer Science: Modelling and Reasoning About Systems*. Cambridge University Press, second ed., 2004.

[7] M. de Rijke, *Extending Modal Logic*. ILLC, 1993.

[8] H. Andréka, I. Németi, and J. van Benthem, "Modal languages and bounded fragments of predicate logic," *Journal of Philosophical Logic*, vol. 27, 1998.

[9] J. R. Burch, E. M. Clarke, K. L. McMillan, D. L. Dill, and L. J. Hwang, "Symbolic model checking: 10^{20} states and beyond," *Proceedings. Fifth Annual IEEE Symposium on Logic in Computer Science*, pp. 428–439, 1990.