Notes on Mathematical Modal Logic

1 Introduction

These notes were written by Wang Yafeng following a course of three intensive lectures on classical themes in mathematical modal logic given by Johan van Benthem in the Berkeley-Stanford Logic Circle, San Francisco, May 2015. Some of this material was collected in the monograph van Benthem [13], [16], other parts come from later publications. We will provide some references to further relevant work in this document, but our bibliography is not self-contained.

Our text starts with a model-based perspective on modal logic. From this perspective, modal logic is just a special fragment of first order logic with certain syntactic restrictions. More precisely, modal logic can be translated into first order logic via the standard translation ST: $M, s \vDash \varphi$ iff $M, s \vDash ST(\varphi)$. For instance, $\Box p$ can be translated as $\forall y(Rxy \to Ry)$.

We then move on to a frame-based perspective on modal logic. A frame is simply a model stripped of its valuation, and the validity of a modal formula in a frame can be translated into second order logic: $F, s \vDash \varphi$ iff $\forall \vec{P} ST(\varphi)$. From this perspective, modal logic is a special fragment of monadic second order logic with all the quantifiers out in front (called Π^1_1 formulas). Interestingly, some modal formulas also correspond to first-order conditions of independent interest, and we would like to understand why they do so.

Many important results in the classical era of modal logic have to do with either of these perspectives. In the first half of these notes, we will mainly prove two results from the model-based perspective: the modal invariance theorem and the modal Lindström theorem. As a highlight from the frame based perspective, we will prove the so-called Sahlqvist Theorem that covers a wide range of modal axioms from the literature, and we will also present some key examples of modal formulas that lack first-order correspondents.

Next we consider a more computational question relevant to our definability concerns: Is there some way to tell in general whether a given modal formula has a first order correspondent? We give a negative answer: it is at least undecidable. We then move on to discuss further definability results, involving ultraproducts of models, which give us a sense of what modal formulas are capable of expressing at the level of models.

The bulk of the second half of the notes is devoted to the Goldblatt-Thomason Theorem, which describes what modal formulas are capable of expressing at the level of frames. Our first proof of this theorem introduces a third important perspective on modal logic, namely the algebraic perspective. We introduce the necessary algebraic tools along the way, construct an algebraic proof of the GT theorem, and then convert it into a second, purely model-theoretic proof. We take this as a case study that leads to further observations on the interplay between model-theoretic and algebraic methods in modal logic. We also provide some further model-theoretic preservation results for the frame constructions involved in the Goldblatt-Thomason theorem.

Finally, as a more recent 'neo-classical' topic, we conclude these notes with a glimpse of two natural logics extending the usual propositional or first-order base logics for modality — infinitary logics and fixed-point logics. These two systems turn out to have natural connections with modal logic at the model or the frame level.

2 Background in First-Order Model Theory

Typically, the proof for the modal invariance theorem rests on a number of first-order concepts and results, all of which are centered on *saturated models* (a standard concept in first-order model theory). This section provides the tools for constructing certain saturated models that are useful in the modal invariance theorem. We start by defining ultrafilters and the ultraproduct construction based on ultrafilters. After that we introduce the notion of countably saturated models, and show how to build countably saturated models by using ultraproducts based on a special kind of ultrafilters—the countably incomplete ultrafilters.

2.1 Ultraproducts

In this section we construct a species of models called *ultraproducts*. Roughly, they are quotients of cartesian produces of models, where the congruence relation is induced by an *ultrafilter* over the index set, which is a special collection of subsets of the index set.

Definition 2.1. ***(Ultrafilters.)*** *Let I be a non-empty set. A **filter** F over I is a set $F \subseteq \mathcal{P}(I)$ such that*

1. *$I \in F$.*

2. *F is closed under finite intersections: If $X, Y \in F$, then $X \cap Y \in F$.*

3. *F is closed under supersets: If $X \in F$ and $X \subseteq Z \subseteq I$, then $Z \in F$.*

*F is **proper** iff $\emptyset \notin F$. An **ultrafilter** over W is a proper filter U such that for all $X \in \mathcal{P}(I)$, $X \in U$ if and only if $(I - X) \notin U$.*

Intuitively, we can think of a filter F over I as a set of all the 'big' subsets of I: The sets in F are so big with respect to I that their finite intersections will always be big, and the superset of a big set is certainly big. An ultrafiler partitions the subsets of I into 'big' and 'small' sets; it leaves out all the small sets and keeps only the big sets. Moreover, it is a maximal filter, in the sense that it cannot be extended further. Not all ultrafilters are created equal, however. In particular, we would like to separate the trivial ultrafilters from the nontrivial ones. An ultrafilter U is trivial (or 'principal') just in case there is an element in the index set I that acts as a 'dictator' of U, and an ultrafilter is nontrivial just in case it is dictator free.

Definition 2.2. *An ultrafilter U over I is **principal** if there is an element $i \in I$ such that for all $X \in \mathcal{P}(I)$, $X \in U$ if and only if $i \in X$.*

Theorem 2.3. ***(Ultrafilter Theorem.)*** *Any proper filter over a non-empty set I can be extended to an ultrafilter over I.*

Proof. (Sketch only.)

Let E be a proper filter over I, and $\mathcal{E} = \{F \mid F \text{ is a proper filter over } I \text{ and } E \subseteq F\}$. It is straightforward to show that if C is a nonempty chain of proper filters in $\mathcal{E}$, then $\bigcup C$ is a proper filter over I and $E \subseteq \bigcup C$; hence $\bigcup C \in \mathcal{E}$. By Zorn's lemma, $\mathcal{E}$ has a maximal element D such that $E \subseteq D$. D is clearly a maximal proper filter over I; because if D' is a proper filter over I such that $D \subseteq D'$, then $E \subseteq D'$, hence $D' \in \mathcal{E}$ and so $D' = D$. But any maximal proper filter over I is also an ultrafilter over I (and vice versa), therefore D is an ultrafilter over I. ■

A set E is said to have the *finite intersection property* iff the intersection of any finite number of elements of E is nonempty. As the following corollary shows, the finite intersection property is very useful for constructing ultrafilters:

Corollary 2.4. *Any subset of $\mathcal{P}(I)$ with the finite intersection property can be extended to an ultrafilter over I.*

Proof. (Sketch only.)

Suppose $C \subseteq \mathcal{P}(I)$ and C has the finite intersection property. Let D be *the filter generated by* C, that is,

$$D = \bigcap \{F \mid C \subseteq F \text{ and } F \text{ is a filter over } I\}$$

Clearly $C \subseteq D$. Moreover, we can easily show that D is a proper filter given the assumption that C has the finite intersection property. By the ultrafilter theorem, D can then be extended to an ultrafilter over I. ■

Suppose U is an ultrafilter over a non-empty set I, and that for each $i \in I$, W_i is a non-empty set. Let $C = \prod_{i \in I} W_i$ be the cartesian product of these sets. Each tuple $f \in \prod_{i \in I} W_i$ can be thought of as a function $f : I \to \bigcup_{i \in I} W_i$ such that for each $i \in I$, $f(i) \in W_i$, with notation $f(i)$ referring to the i-th coordinate of the tuple. We can then define a relation $\sim_U$ on C as follows: For $f, g \in C$, $f \sim_U g$ iff $\{i \in I \mid f(i) = g(i)\} \in U$. Intuitively, $\sim_U$ relates two tuples f and g if and only if the set of indices of coordinates on which f and g agree is a 'big' set (i.e., in the ultrafilter U), and we say that the tuples agree on 'sufficiently many' or 'U-many' coordinates. Using the fact that U is a filter, it is straightforward show that $\sim_U$ is an equivalence relation on C; hence we can define the equivalence class $[f]_U = \{g \in C \mid g \sim_U f\}$ for every $f \in C$.

We are now ready to introduce the ultraproduct construction for models:

Definition 2.5. ***(Ultraproducts.)*** *Let $\mathcal{L}$ be a first order language, $\{\mathfrak{M}_i\}_{i \in I}$ be a family of $\mathcal{L}$-models, and U be an ultrafilter over I. For every $f, g \in \prod_{i \in I} W_i$ (W_i is the universe of $\mathfrak{M}_i$), define*

$f \sim_U g$ iff $\{i \in I \mid f(i) = g(i)\} \in U$, and denote the equivalence class of f under $\sim_U$ as $[f]_U$.

*The **ultraproduct** $\prod_U \mathfrak{M}_i$ of $\mathfrak{M}_i$ modulo U is the model defined as:*

1. *The universe W_U of $\prod_U \mathfrak{M}_i$ is the quotient $\prod_{i \in I} W_i / \sim_U$, that is: the collection of equivalence classes $\{[f]_U \mid f \in \prod_{i \in I} W_i\}$.*

2. *Let R be any n-place relation symbol, and $R^{\mathfrak{M}_i}$ its interpretation in the model $\mathfrak{M}_i$. The interpretation R^U of R in $\prod_U \mathfrak{M}_i$ is given by*

$$R^U([f_1]_U, \dots, [f_n]_U) \text{ iff } \{i \in I \mid R^{\mathfrak{M}_i}(f_1(i), \dots, f_n(i))\} \in U$$

3. *Let F be any n-place function symbol, and $F^{\mathfrak{M}_i}$ its interpretation in the model $\mathfrak{M}_i$. The interpretation F^U of F in $\prod_U \mathfrak{M}_i$ is given by*

$$F^U([f_1]_U, \dots, [f_n]_U) = [\lambda i. F^{\mathfrak{M}_i}(f_1(i), \dots, f_n(i))]_U$$

   Where $\lambda i. F^{\mathfrak{M}_i}(f_1(i), \dots, f_n(i))$ is the tuple in $\prod_{i \in I} W_i$ whose i-th coordinate is $F^{\mathfrak{M}_i}(f_1(i), \dots, f_n(i))$.

4. *Let c be any constant symbol, and $c^{\mathfrak{M}_i}$ its interpretation in the model $\mathfrak{M}_i$. The interpretation c^U of c in $\prod_U \mathfrak{M}_i$ is given by*

$$c^U = [\lambda i. c^{\mathfrak{M}_i}]_U$$

   Where $\lambda i. c^{\mathfrak{M}_i}$ is the tuple in $\prod_{i \in I} W_i$ whose i-th coordinate is $c^{\mathfrak{M}_i}$.

*In the case where $\mathfrak{M}_i = \mathfrak{M}$ for all $i \in I$, we say that $\prod_U \mathfrak{M}$ is the **ultrapower** of $\mathfrak{M}$ modulo U.*

It is straightforward to check that the functions and relations are well defined; i.e., whether $R^U[f_1]_U \dots [f_n]_U$ holds, and what the image $F^U([f_1]_U, \dots, [f_n]_U)$ is, do not depend on the choice of representatives $f_1, \dots, f_n$ for the equivalence classes $[f_1]_U, \dots, [f_n]_U$. Moreover, these verifications depend only on the closure of filters under supersets and finite intersections; i.e., $\prod_U \mathfrak{M}_i$ is well-defined even if U is a filter rather than an ultrafilter.

Now we introduce the theorem of crucial importance for the study of ultraproducts:

Theorem 2.6. ***(The Fundamental Theorem of Ultraproducts.)*** *Let $\mathfrak{N}$ be the ultraproduct $\prod_U \mathfrak{M}_i$, and let I be the index set. Then:*

1. *For any term $t(x_1 \dots x_n)$ of $\mathcal{L}$ and elements $[f_1]_U, \dots, [f_n]_U \in \mathfrak{N}$, we have*

$$t^{\mathfrak{N}}([f_1]_U, \dots, [f_n]_U) = [\lambda i.t^{\mathfrak{M}_i}(f_1(i) \dots f_n(i))]_U$$

2. *Given any formula $\varphi(x_1 \dots x_n)$ of $\mathcal{L}$ and $[f_1]_U, \dots, [f_n]_U \in \mathfrak{N}$, we have*

$$\mathfrak{N} \vDash \varphi[[f_1]_U, \dots, [f_n]_U] \quad \textit{iff} \quad \{i \in I \mid \mathfrak{M}_i \vDash \varphi[f_1(i) \dots f_n(i)]\} \in U$$

Proof. (Sketch only.)

For part (2) we use induction on formulas. That the biconditional holds for atomic formulas follows straightforwardly from the definition of ultraproducts. The induction step for Boolean connectives is equally straightforward. (Note: the case for negation is the *only* place that uses the assumption that U is an *ultra*filter.) To prove the induction step for the $\exists$-case, let φ be of the form $\exists x \psi(x)$, and assume the biconditional holds for ψ, where x is free in ψ. We show that it also holds for $\exists x \psi(x)$.

($\Longrightarrow$): Suppose $\mathfrak{N} \vDash \exists x \psi(x)$. This means that there exists $[f]_U \in \mathfrak{N}$ such that $\mathfrak{N} \vDash \psi[[f]_U]$. By our inductive hypothesis, this means $\{i \in I \mid \mathfrak{M}_i \vDash \psi[f(i)]\} \in U$. Since $\{i \in I \mid \mathfrak{M}_i \vDash \psi[f(i)]\} \subseteq \{i \in I \mid \mathfrak{M}_i \vDash \exists x \psi\}$, we have that $\{i \in I \mid \mathfrak{M}_i \vDash \exists x \psi\} \in U$ by the $\subseteq$-closure of U.

($\Longleftarrow$): Suppose $\{i \in I \mid \mathfrak{M}_i \vDash \exists x \psi\} \in U$. Then, for each $i \in \{i \in I \mid \mathfrak{M}_i \vDash \exists x \psi\}$, we use the axiom of choice to pick an element $a_i \in \mathfrak{M}_i$ such that $\mathfrak{M}_i \vDash \psi[a_i]$; and for each $i \notin \{i \in I \mid \mathfrak{M}_i \vDash \exists x \psi\}$, we randomly pick an element $a_i \in M_i$. Take the tuple $f := \lambda i.a_i$. Since $f(i) = a_i$ for all i, We clearly have $\{i \in I \mid \mathfrak{M}_i \vDash \exists x \psi\} \subseteq \{i \in I \mid \mathfrak{M}_i \vDash \psi[f(i)]\}$ by our construction. So, by $\subseteq$-closure of U, we have $\{i \in I \mid \mathfrak{M}_i \vDash \psi[f(i)]\} \in U$. By inductive hypothesis, $\mathfrak{N} \vDash \psi[[f]_U]$. This entails $\mathfrak{N} \vDash \exists x \psi$, and we are done. ■

Intuitively, the fundamental theorem of ultraproducts says that ultraproducts preserve exactly those first order formulas which hold in sufficiently many (U-many) models $\mathfrak{M}_i$. Chris Mierzewski [12] suggests a nice metaphor which conceives of taking ultraproducts as a voting scenario: We can think of an ultrafilter U on I as a collection of 'parties' or 'coalitions' that have a significant voice. Each model votes for the first-order properties it possesses, and the fundamental theorem of ultraproducts says that a first order property is agreed on in the final vote (the ultraproduct) exactly when a 'significant' coalition unanimously agreed on it.

Definition 2.7. *Let $\mathfrak{M}$ and $\mathfrak{N}$ be two $\mathcal{L}$-models. An **elementary embedding** h of $\mathfrak{M}$ into $\mathfrak{N}$ is an injection that preserves all the first order formulas. That is, for every $\varphi \in \mathcal{L}$ and $a_1, \dots a_n \in \mathfrak{M}$,*

$$\mathfrak{M} \models \varphi[a_1, \dots, a_n] \text{ iff } \mathfrak{N} \models \varphi[h(a_1), \dots, h(a_n)]$$

Corollary 2.8. *Let $\prod_U \mathfrak{M}$ be an ultrapower of $\mathfrak{M}$. Define the **canonical embedding** $d : \mathfrak{M} \to \prod_U \mathfrak{M}$ as:*

$$d(a) = [f_a]_U, \text{ where } f_a(i) = a, \text{ for all } i \in I.$$

Then d is an elementary embedding.

Proof. d is clearly an injection. Moreover, let $\varphi(x_1, \dots, x_n)$ be any formula of $\mathcal{L}$ and let $a_0, \dots, a_n$ be any elements in the universe of $\mathfrak{M}$. Then, by the fundamental theorem of ultraproducts,

$$\begin{aligned} \prod_U \mathfrak{M} \models \varphi[d(a_0), \dots, d(a_n)] \quad &\text{iff} \quad \prod_U \mathfrak{M} \models \varphi[[f_{a_0}]_U, \dots, [f_{a_n}]_U] \\ &\text{iff} \quad \{i \in I \mid \mathfrak{M} \models \varphi[f_{a_0}(i), \dots, f_{a_n}(i)]\} \in U \\ &\text{iff} \quad \{i \in I \mid \mathfrak{M} \models \varphi[a_0, \dots, a_n]\} \in U \\ &\text{iff} \quad \mathfrak{M} \models \varphi[a_0, \dots, a_n] \end{aligned}$$

Therefore d is an elementary embedding. ■

The fundamental theorem of ultraproducts gives us a very natural proof of the following major theorem of model theory, the Compactness Theorem for first order logic.

Theorem 2.9. ***(Compactness Theorem.)*** *A set of sentences Σ of a first order language $\mathcal{L}$ has a model if and only if every finite subset of Σ has a model.*

Proof. The left-to-right direction is trivial; we prove the other. Let the index set I be the set of all finite subsets of Σ, and assume that each $\Delta \in I$ has a model $\mathfrak{M}_\Delta$. For each $\Delta \in I$ define $S_\Delta = \{\Gamma \in I \mid \Delta \subseteq \Gamma\}$. Consider the set $S = \{S_\Delta \mid \Delta \in I\}$: it clearly has the finite intersection property, since given S_Δ and S_Γ in S, we have $S_\Delta \cap S_\Gamma = S_{\Delta \cap \Gamma} \neq \emptyset$. By corollary 2.4, there is an ultrafilter U on I such that $S \subseteq U$.

Now take the ultraproduct $\prod_U \mathfrak{M}_\Delta := \prod_{\Delta \in I} \mathfrak{M}_\Delta / \sim_U$. We claim that this is a model of Σ: Take any $\varphi \in \Sigma$. We have $S_{\{\varphi\}} \in U$ and also $S_{\{\varphi\}} \subseteq \{\Gamma \in I \mid \mathfrak{M}_\Gamma \models \varphi\}$, so by $\subseteq$-closure of ultrafilters, we have $\{\Gamma \in I \mid \mathfrak{M}_\Gamma \models \varphi\} \in U$. The fundamental theorem of ultraproducts implies $\prod_U \mathfrak{M}_\Delta \models \varphi$, as desired. ■

2.2 Saturation

Saturated models are models rich enough to realize all the complete descriptions (of a point) which are consistent with the first-order theory of the model. Basically, saturation means big enough to include all the logically consistent potential points. To make our description powerful enough, we may introduce new constants to talk about the points in the given model directly. Saturation is usually defined via complete 1-types (complete FO-descriptions of a potential point), here we use an alternative definition. The equivalence of the definitions is based on: Σ is a type with respect to a model $\mathfrak{M}$ if and only if Σ is consistent with $\mathrm{Th}(\mathfrak{M})$ if and only if Σ is finitely satisfiable in $\mathfrak{M}$.

To introduce some notation, take a first order language $\mathcal{L}$ and a $\mathcal{L}$-model $\mathfrak{M}$ with domain W. For a subset $A \subseteq W$, $\mathcal{L}[A] = \mathcal{L} \cup \{\bar{a} \mid a \in A\}$ is the language obtained by extending $\mathcal{L}$ with new constants $\bar{a}$ for all elements $a \in A$. Let $\mathfrak{M}_A = (\mathfrak{M}, a)_{a \in A}$ be the expansion of $\mathfrak{M}$ to a structure for $\mathcal{L}[A]$ in which each $\bar{a}$ is interpreted as a, and let $\mathrm{Th}(\mathfrak{M}_A)$ be the set of all $\mathcal{L}[A]$-sentences that are true in $\mathfrak{M}_A$.

Definition 2.10. ***(Saturated Models.)*** *Let α be a cardinal number, $\mathcal{L}$ be a first order language and $\mathfrak{M}$ be a $\mathcal{L}$-model with domain W. $\mathfrak{M}$ is **α-saturated** if for every subset $A \subseteq W$ of size less than α, and for every set $\Sigma(x)$ of $\mathcal{L}[A]$-formulas in the free variable x, if $\Sigma(x)$ is finitely satisfiable in $\mathfrak{M}_A$ then $\Sigma(x)$ is satisfiable in $\mathfrak{M}_A$. An ω-saturated model is usually called **countably saturated**.*

When working with α-saturated models, we frequently want to enumerate the elements of A in the expansion $(\mathfrak{M}, a)_{a \in A}$. Thus if $\xi < \alpha$ and $A = \{a_\eta : \eta < \xi\}$, we write $(\mathfrak{M}, a_\eta)_{\eta < \xi}$.

To build countably saturated models, we use ultrapowers based on a special kind of ultrafilter that fails to be closed under countable intersections.

Definition 2.11. *A filter D is **countably incomplete** iff it is not closed under countable intersections; that is, there is a countable set $E \subseteq D$ such that $\bigcap E \notin D$.*

It is not hard to see that every principal ultrafilter is not countably incomplete, so a countably incomplete ultrafilter must be non-principal. Such ultrafilters exist. For instance, the set of all co-finite subsets of $\mathbb{N}$ has the intersection property, and so it can be extended to an ultrafilter U. Then, for all $n \in \mathbb{N}$, $\mathbb{N} \backslash \{n\} \in U$. But $\bigcap_{n \in \mathbb{N}} (\mathbb{N} \backslash \{n\}) = \emptyset \notin U$, hence U is countably incomplete. However, note that the existence of uncountable cardinals that admit a non-principal countably incomplete ultrafilter (called measurable cardinals) is not provable in ZFC. (cf. [10])

Lemma 2.12. *Let U be a countably incomplete ultrafilter over the set I. Then there exists a countable decreasing chain*

$$I = I_0 \supseteq I_1 \supseteq I_2 \ldots$$

of elements $I_n \in U$ such that $\bigcap_n I_n = \emptyset$.

Proof. Let U be a countably incomplete ultrafilter over I. Take a set $E = \{e_0, e_1, \ldots\} \subseteq U$ such that $\bigcap E \notin U$, so $I - \bigcap E \in U$ since U is an ultrafilter. Define a countable decreasing chain $E' = \{e'_0, e'_1, \ldots\}$ such that

$$e'_0 = e_0 \cap (I - \cap E)$$

$$e'_{n+1} = e'_n \cap e_{n+1}$$

Clearly $E' \subseteq U$, and $\bigcap E' = (\bigcap E) \cap (I - \bigcap E) = \emptyset$ as desired. ■

Lemma 2.13. *Let $\mathcal{L}$ be a countable language, $\mathfrak{M}_i$ $(i \in I)$ be a family of $\mathcal{L}$ models, U be a countably incomplete ultrafilter over the set I, and $\prod_U \mathfrak{M}_i$ be the ultraproduct of $\mathfrak{M}_i$ $(i \in I)$. For every set $\Sigma(x)$ of formulas of $\mathcal{L}$, if every finite subset of $\Sigma(x)$ is satisfiable in $\prod_U \mathfrak{M}_i$, then $\Sigma(x)$ is satisfiable in $\prod_U \mathfrak{M}_i$.*

Proof. Suppose every finite subset of $\Sigma(x)$ is satisfiable in $\prod_U \mathfrak{M}_i$. Since $\mathcal{L}$ is countable, $\Sigma(x)$ is countable, and we can write

$$\Sigma(x) = \{\sigma_1(x), \sigma_2(x), \ldots, \sigma_n(x), \ldots\}$$

Since U is countably incomplete, by lemma 2.12, there is a countable decreasing chain

$$I = I_0 \supseteq I_1 \supseteq I_2 \ldots$$

such that each $I_n \in U$ and $\bigcap_{n<\omega} I_n = \emptyset$.

Let $X_0 = I$ and for each positive $n < \omega$, define

$$X_n = I_n \cap \{i \in I \mid \mathfrak{M}_i \models (\exists x)(\sigma_1(x) \wedge \dots \sigma_n(x))\}$$

We make two observations about X_n: First, for all $n < \omega$, $X_n \in U$. To see this, note that since $\Sigma(x)$ is finitely satisfiable in $\prod_U \mathfrak{M}_i$, $\sigma_1(x) \wedge \dots \sigma_n(x)$ is satisfiable in $\prod_U \mathfrak{M}_i$. Hence $\prod_U \mathfrak{M}_i \models (\exists x)(\sigma_1(x) \wedge \dots \sigma_n(x))$ and, by the fundamental theorem of ultraproducts, $\{i \in I \mid \mathfrak{M}_i \models (\exists x)(\sigma_1(x) \wedge \dots \sigma_n(x)\} \in U$. Since $I_n \in U$ and U is closed under finite intersections, $X_n \in U$.

Second, for each $i \in I$, there is a greatest number $\max(i) < \omega$ such that $i \in X_{\max(i)}$. This is simply because $\bigcap_{n<\omega} X_n = \emptyset$ and $X_n \supseteq X_{n+1}$.

Now we construct an element $[f]_U \in \prod_U \mathfrak{M}_i$ such that $[f]_U$ satisfies $\Sigma(x)$ in $\prod_U \mathfrak{M}_i$. Let W_i be the domain of $\mathfrak{M}_i$. We define $f \in \prod_{i \in I} W_i$ as follows:

$$\begin{cases} \text{choose } f(i) \text{ to be an arbitrary element of } W_i & \text{if } \max(i) = 0, \\ \text{choose } f(i) \in W_i \text{ so that } \mathfrak{M}_i \models \sigma_1 \wedge \dots \sigma_{\max(i)}[f(i)] & \text{if } \max(i) > 0. \end{cases}$$

Take $\sigma_n(x) \in \Sigma(x)$ for any $n \geq 1$ (note that our enumeration of $\Sigma(x)$ starts with index 1), we want to show that $\prod_U \mathfrak{M}_i \models \sigma_n(x)[[f]_U]$. By the $\subseteq$-closure of U and by the fundamental theorem of ultraproducts, it suffices to show that $X_n \subseteq \{i \in I \mid \mathfrak{M}_i \models \sigma_n[f(i)]\}$. Suppose $i \in X_n$; we have that $\mathfrak{M}_i \models (\exists x)(\sigma_1(x) \wedge \dots \sigma_n(x))$, and that $1 \leq n \leq \max(i)$. But the selection of $f(i)$ ensures that $\mathfrak{M}_i \models \sigma_1 \wedge \dots \sigma_n \dots \wedge \sigma_{\max(i)}[f(i)]$, which implies $\mathfrak{M}_i \models \sigma_n[f(i)]$. Hence $X_n \subseteq \{i \in I \mid \mathfrak{M}_i \models \sigma_n[f(i)]\}$ as desired. ■

Theorem 2.14. *Let $\mathcal{L}$ be a countable first order language, U be a countably incomplete ultrafilter over a non-empty set I, and $\mathfrak{M}$ be a $\mathcal{L}$ model. The ultrapower $\prod_U \mathfrak{M}$ is countably saturated.*

Proof. We show that $\prod_U \mathfrak{M}$ is ω_1−saturated (which implies that $\prod_U \mathfrak{M}$ is ω−saturated). Let $[a_0]_U, \dots, [a_m]_U \dots$ $(m < \omega)$ be a countable sequence of elements of $\prod_U \mathfrak{M}$, $\mathcal{L} \cup \{c_0, c_1, \dots\}$ be the expanded language where c_m is a new constant for $[a_m]_U$, and $\Sigma(x)$ be any set of formulas of $\mathcal{L} \cup \{c_0, c_1, \dots\}$. We want to show that if every finite subset of $\Sigma(x)$ is satisfiable in $(\prod_U \mathfrak{M}, [a_m]_U)_{m<\omega}$, then Σ is satisfiable in $(\prod_U \mathfrak{M}, [a_m]_U)_{m<\omega}$. Note that since $[a_m]_U = \langle a_m(i) : i \in I\rangle_U$,

$$(\prod_U \mathfrak{M}, [a_m]_U)_{m<\omega} = \prod_U ((\mathfrak{M}, a_m(i))_{m<\omega})$$

That is, the expansion $(\prod_U \mathfrak{M}, [a_m]_U)_{m<\omega}$ is itself an ultraproduct of models for the expanded language $\mathcal{L} \cup \{c_0, c_1, \dots\}$. Since $\mathcal{L}$ is a countable language and $\mathcal{L} \cup \{c_0, c_1, \dots\}$ is also countable, what we want to prove follows directly from Lemma 2.13. ■

Corollary 2.15. *Any model for a countable first order language can be elementarily embedded into a countably saturated model.*

Proof. This is a direct consequence of corollary 2.8 and theorem 2.14: Just use the canonical embedding of $\mathfrak{M}$ into the ultrapower $\prod_U \mathfrak{M}$ for some countably incomplete ultrafilter U. ■

Corollary 2.15 can be generalized, although the generalizations will involve heavier model-theoretic machinery that will not be discussed here. We will only mention a generalization of Corollary 2.15 needed later in the proof of the Goldblatt-Thomason Theorem:

Theorem 2.16. *Let α be an infinite cardinal and let $\mathcal{L}$ be a first order language with cardinality α. Then any model for $\mathcal{L}$ can be elementarily embedded into a countably saturated model.*

Proof. This follows from Theorem 6.1.4 and Theorem 6.1.8 of Chang and Keisler [6]. ■

3 Modal Invariance Theorem

Corollary 2.15 enables us to prove a 'bisimulation somewhere else' result: modal equivalence implies bisimulation in some suitably related, countably saturated ultrapowers. This 'bisimulation somewhere else' result in turn provides us a first proof of the modal invariance theorem.

After this proof, we discuss a second proof of the modal invariance theorem by Andréka et al in their paper 'Modal Languages and Bounded Fragments of Predicate Logic'. The second proof also identifies a key lemma of 'transfer' between modal and classical reasoning, although the key lemma is no longer 'bisimulation somewhere else' but rather 'elementary equivalence somewhere else'.

Finally, we note that a third quite different proof will follow our later proof of the modal Lindström Theorem, while one suggestive reformulation will be found in the section on interpolation theorems at the end of these notes.

3.1 The First Proof

Definition 3.1. ***(Bisimulations.)*** *Let $\mathfrak{M} = (W, R, V)$ and $\mathfrak{M}' = (W', R', V')$ be two models for the basic modal language. A non-empty relation $Z \subseteq W \times W'$ is a* ***bisimulation*** *between $\mathfrak{M}$ and $\mathfrak{M}'$ just in case:*

1. *(Atomic Harmony.) If $(x, x') \in Z$, then x and x' satisfies the same proposition letters.*
2. *(Zig.) If $(x, x') \in Z$ and Rxy, then there exists $y' \in W'$ such that $R'x'y'$ and $(y, y') \in Z$.*
3. *(Zag.) If $(x, x') \in Z$ and $R'x'y'$, then there exists $y \in W$ such that Rxy and $(y, y') \in Z$.*

Z is a bisimulation between two pointed models $(\mathfrak{M}, w)$ and $(\mathfrak{M}', w')$ if Z is a bisimulation between $\mathfrak{M}$ and $\mathfrak{M}'$, and $(w, w') \in Z$. (We often use $\underline{\leftrightarrow}$ as the notation for bisimulation.)

Lemma 3.2. *Suppose $(\mathfrak{M}, w)$ and $(\mathfrak{N}, v)$ are modally equivalent. Then $\mathfrak{M}$ and $\mathfrak{N}$ can be elementarily embedded into $\omega-$saturated models $\mathfrak{M}^+$ and $\mathfrak{N}^+$ respectively, such that there exists a bisimulation $\underline{\leftrightarrow}$ between $(\mathfrak{M}^+, w)$ and $(\mathfrak{N}^+, v)$.*

Proof. Suppose $\mathfrak{M}, w \vDash \varphi$ iff $\mathfrak{N}, v \vDash \varphi$. By corollary 2.15, there exist $\omega-$saturated elementary extensions $\mathfrak{M}^+$, $\mathfrak{N}^+$ of $\mathfrak{M}$, $\mathfrak{N}$ respectively. Note that $\mathfrak{M}^+, w$ and $\mathfrak{N}^+, v$ are modally equivalent: for any modal formula φ,

$$\begin{aligned} \mathfrak{M}^+, w \vDash \varphi \ \text{ iff } \ & \mathfrak{M}, w \vDash \varphi && \text{(elementary embedding)} \\ \text{iff } \ & \mathfrak{N}, v \vDash \varphi && \text{(modal equivalence)} \\ \text{iff } \ & \mathfrak{N}^+, v \vDash \varphi && \text{(elementary embedding)} \end{aligned}$$

Now we prove that, in these $\omega-$saturated models, the above-defined relation of modal equivalence is itself a bisimulation.

(Atomic Harmony): If $\mathfrak{M}^+, w$ and $\mathfrak{N}^+, v$ are modally equivalent, then in particular they verify all the proposition letters.

(Zig): Suppose $\mathfrak{M}^+, s$ and $\mathfrak{N}^+, t$ are modally equivalent and $R^{\mathfrak{M}^+}sw$. We want to show that there exists v in $\mathfrak{N}^+$ such that $R^{\mathfrak{N}^+}tv$ and that $\mathfrak{M}^+, w$ and $\mathfrak{N}^+, v$ are modally equivalent. Let $T(x)$ be the set of all (standard translations of) modal formulas true at w in $\mathfrak{M}^+$:

$$T(x) = \{ST_x(\psi) \mid \mathfrak{M}^+, w \models ST_x(\psi))\}$$

and let $\bar{t}$ be a constant for t. We show that the following set of formulas $\{R\bar{t}x\} \cup T(x)$ is finitely satisfiable in the expansion $(\mathfrak{N}^+, t)$. Suppose $T_0(x)$ is an arbitrary finite subset of $T(x)$, and we have $\mathfrak{M}^+, w \models \bigwedge T_0(x)$. Hence $\mathfrak{M}^+, s \models \exists x(Ryx \wedge \bigwedge T_0(x))$, and $\exists x(Ryx \wedge \bigwedge T_0(x))$ is the standard translation of a modal formula. Given the assumption that $\mathfrak{M}^+, s$ and $\mathfrak{N}^+, t$ are modally equivalent, we have $\mathfrak{N}^+, t \models \exists x(Ryx \wedge \bigwedge T_0(x))$. Hence, there exists v_0 in $\mathfrak{N}^+$ such that $\mathfrak{N}^+, v_0 \models R\bar{t}x \wedge \bigwedge T_0(x)$.

Now, the fact that $\mathfrak{N}^+$ is $\omega-$saturated implies that the set of formulas $\{R\bar{t}x\} \cup T(x)$ is satisfiable in $\mathfrak{N}^+$. Hence there exists v in $\mathfrak{N}^+$ such that $R^{\mathfrak{N}^+}tv$ and that the entire modal theory $T(x)$ of w is true at v—that is, $\mathfrak{M}^+, w$ and $\mathfrak{N}^+, v$ are modally equivalent.

(Zag): Similar proof. ■

When proving Lemma 3.2, we prove the claim that the relation of modal equivalence in countably saturated models is a bisimulation. This claim can be strengthened, however. (The proof of Lemma 3.2, for instance, only requires 2-saturation to go through.) We can introduce a weaker notion of *modally saturated models*, and show that the relation of modal equivalence in modally saturated models is a bisimulation.

Definition 3.3. ***(Modal Saturation)***. *Let $\mathfrak{M} = (W, R, V)$ be a modal model, X a subset of W and Σ a set of modal formulas. Σ is satisfiable in the set X if there is a point $x \in X$ such that $\mathfrak{M}, x \models \varphi$ for all $\varphi \in \Sigma$. Σ is finitely satisfiable in X if every finite subset of Σ is satisfiable in X. We call the model $\mathfrak{M}$* ***modally saturated*** *if for every point $w \in W$ and every set Σ of modal formulas, if Σ is finitely satisfiable in the set of successors of w, then Σ is satisfiable in the set of successors of w.*

Lemma 3.4. *Any countably saturated model is modally saturated.*

Lemma 3.5. *Let $\mathfrak{M}$ and $\mathfrak{M}'$ be two modally saturated models. Then the relation of modal equivalence between points in $\mathfrak{M}$ and points in $\mathfrak{M}'$ is a bisimulation.*

Lemma 3.4 and 3.5 can be easily proved using the ideas in the proof of Lemma 3.2, and they will be useful later on when we prove the Goldblatt-Thomason Theorem. For now, let us go back and finish the first proof of the Modal Invariance Theorem:

Theorem 3.6. ***(Modal Invariance Theorem).*** *Let $\varphi = \varphi(x)$ be a formula (with one free variable) in the first order language of modal models. The following two assertions are equivalent: (a) φ is logically equivalent to (the standard translation of) a modal formula, (b) φ is invariant for bisimulation.*

Proof. $(a \Rightarrow b)$ is done by induction on modal formulas.

$(a \Leftarrow b)$ Assume that $\varphi(x)$ is invariant for bisimulation. Let $\mathrm{Mod}(\varphi)$ be the set of modal consequences of φ:

$$\mathrm{Mod}(\varphi) = \{ST_x(\psi) \mid \psi \text{ is a modal formula, and } \varphi(x) \vDash ST_x(\psi)\}$$

If we can show that $\mathrm{Mod}(\varphi) \vDash \varphi$, then we can show that φ is equivalent to (the standard translation of) a modal formula. To see why, suppose $\mathrm{Mod}(\varphi) \vDash \varphi$. By the Compactness for first order logic, there exists some finite subset $X \subseteq \mathrm{Mod}(\varphi)$ such that $X \vDash \varphi$, and so $\bigwedge X \vDash \varphi$. We assume that $\varphi \vDash \bigwedge X$, thus φ is equivalent to $\bigwedge X$, which is the translation of a modal formula.

Now let us prove the claim that $\mathrm{Mod}(\varphi) \vDash \varphi$. Assume $\mathfrak{M}, w \vDash \mathrm{Mod}(\varphi)$, we want to show that $\mathfrak{M}, w \vDash \varphi(x)$. Let $T(x)$ be the set of all (standard translations of) modal formulas true at w in $\mathfrak{M}$:

$$T(x) = \{ST_x(\psi) \mid \mathfrak{M}, w \vDash ST_x(\psi))\}$$

It is easy to see that $T(x) \cup \{\varphi(x)\}$ is finitely satisfiable: If not, then there exists a finite subset $T_0(x) \subseteq T(x)$ such that $\varphi(x) \vDash \neg \bigwedge T_0(x)$. Hence $\neg \bigwedge T_0(x) \in \mathrm{Mod}(\varphi)$. But this implies that $\mathfrak{M}, w \vDash \neg \bigwedge T_0(x)$, contradicting our assumption that $T_0(x) \subseteq T(x)$ and that $\mathfrak{M}, w \vDash T(x)$.

By compactness for first order logic, there exists $\mathfrak{N}, v$ such that $\mathfrak{N}, v \vDash T(x) \cup \{\varphi(x)\}$. Since the entire modal theory $T(x)$ of w is true at v, w and v are modally equivalent: for all modal formulas ψ, $\mathfrak{M}, w \vDash \psi$ iff $\mathfrak{N}, v \vDash \psi$. Now, by lemma 3.2, there exist $\omega-$saturated elementary extensions $\mathfrak{M}^+$, $\mathfrak{N}^+$ of $\mathfrak{M}$, $\mathfrak{N}$ respectively, such that $\mathfrak{M}^+, w \underline{\leftrightarrow} \mathfrak{N}^+, v$:

$$\begin{tikzcd}
\mathfrak{M}, w \arrow[r, "\equiv"'] & \mathfrak{N}, v \arrow[d, "\preceq"] \\
\mathfrak{M}^+, w \arrow[u, "\preceq"] & \mathfrak{N}^+, v \arrow[l, "\underline{\leftrightarrow}"]
\end{tikzcd}$$

Since $\mathfrak{N}, v \vDash \varphi(x)$ and the truth of first-order formulas is preserved under elementary embeddings, $\mathfrak{N}^+, v \vDash \varphi(x)$. As $\varphi(x)$ is invariant under bisimulation, $\mathfrak{M}^+, w \vDash \varphi(x)$. Again by invariance under elementary embedddings, we have $\mathfrak{M}, w \vDash \varphi(x)$ as desired. ■

3.2 The Second Proof.

Andréka et al [1] provides another proof of the modal invariance theorem. The basic idea of this proof is that modal equivalence can be 'updated' to full elementary equivalence up to bisimulation:

Lemma 3.7. *Suppose two models $\mathfrak{M}, w$ and $\mathfrak{N}, v$ are modally equivalent. Then they possess bisimulations with two models $\mathfrak{M}^*, w$ and $\mathfrak{N}^*, v$ respectively which are elementarily equivalent.*

Proof. The required models $\mathfrak{M}^*, w$ and $\mathfrak{N}^*, v$ are the tree unraveling of $\mathfrak{M}, w$ and $\mathfrak{N}, v$ with modifications. For instance, the domain of $\mathfrak{M}^*, w$ consists of finite sequences of the form $u = (w, u_1, \ldots, u_k)$, where $wR^{\mathfrak{M}}u_1$ and each u_{i+1} is a $R^{\mathfrak{M}}$ successor of u_i in $\mathfrak{M}$. $(w, u_1, \ldots, u_n)R^{\mathfrak{M}^*}(w, v_1, \ldots, v_m)$ just in case $m = n + 1$, $u_i = v_i$ for $i = 1, \ldots, n$, and $u_n R^{\mathfrak{M}} v_m$. The valuation V^* is defined so that $(w, u_1, \ldots, u_n) \in V^*(p)$ iff $u_n \in V(p)$. Bisimulation with the original model is obtained by connecting each sequence with its last element.

For what follows, in addition to the preceding standard unraveling, we also perform *multiplication*, which makes sure that each node except the root gets copied infinitely many times. This can be done as follows: First, copy each successor of w at level 1 countably many times, and attach these disjoint copies to w. Next, consider successors at level 2 on all branches of the previous stage, and perform the same copying process. At each stage, there is an obvious bisimulation (connecting copies with originals). Iterating this process through all finite levels yields the intended model $\mathfrak{M}^*, w$, and similarly, we obtain a 'multiplied unraveling' $\mathfrak{N}^*, v$.

We show that $\mathfrak{M}^*, w$ and $\mathfrak{N}^*, v$ are elementary equivalent by using Ehrenfeucht Games. It suffices to show that the Duplicator can win a game of n rounds for any finite n; that is, there is a partial isomorphism between the two structures after n rounds.

In analyzing this game, we know that the two roots w and v in the unraveled multiplied models have the same modal theory. In fact, it follows from this that they also have the same tense-logical theory (in the basic modal language extended with a backward modal operator for 'past'). This observation helps us define the partial isomorphism between the two models.

Suppose that in round i of the game, a match $a \equiv b$ has been established between a finite number of worlds a, b in the two models which satisfy the following three conditions:

1. If $a \equiv b$, then $(\mathfrak{M}^*, a)$ is equivalent with $(\mathfrak{N}^*, b)$
   for all the tense-logical formulas up to degree 2^{n-i}.

2. If $a \equiv b$ and $a' \equiv b'$, and the distance between a and a' in $\mathfrak{M}^*$ is at most 2^{n-i},
   then the distance between b and b' in $\mathfrak{N}^*$ is the same, and also, the path between
   a and a' is isomorphic to the path between b and b' via $\equiv$. (Here computing
   'distance' between points may include backtracking along the tree, which is why
   we will use two-sided tense-logical formulas to describe the paths.)

3. If the distance between the distance between a and a' in $\mathfrak{M}^*$ is greater than 2^{n-i},
   then the distance between b and b' in $\mathfrak{N}^*$ is greater than 2^{n-i} as well.

We show that no matter what Spoiler chooses, Duplicator can maintain the matching in the next step. Suppose Spoiler's next choice is some point P in either tree. There are two cases:

Case 1. P has distance $\leq 2^{N-i-1}$ to some point Q that was already matched at the previous stage, say to some point Q' in the other model. Consider the unique path of length $k \leq 2^{N-i-1}$ between P and Q, and attach complete tense-logical descriptions δ to its nodes up to degree 2^{N-i-1} (in particular, δ_k is the complete tense-logical description of P up to degree 2^{N-i-1}). Then the following tense-logical formula that describes the path is true at point Q in the tree:

$$\mathrm{PAST}(\delta_1 \wedge \mathrm{PAST}(\cdots \wedge \mathrm{PAST}(\delta_i \wedge \mathrm{FUT}(\delta_{i+1} \wedge ... \wedge \mathrm{FUT}(\delta_k))$$

The total degree of this formula is at most 2^{N-i-1} (the degree of the descriptions of the nodes) + 2^{N-i-1} (the length of the path) = 2^{N-i}. Since in round i, Q and Q' agree on tense-logical formulas up to degree 2^{N-i}, this formula is true at Q' as well. Hence we can find corresponding points in the other model, making the two paths isomorphic while also maintaining tense-logical equivalence up to 2^{N-i-1} at P and its matching point P'.

Case 2. P has distance $> 2^{N-i-1}$ from all previously matched points. Take the unique path from the root (w or v) to P, and again describe the nodes of this path with tense-logical descriptions up to degree 2^{N-i-1}, and then describe the entire path in a tense-logical formula Δ. (This time there is no syntactic depth restriction on the total formula.) Since the two roots agree on *all* tense logical formulas, Δ is true at the other root as well, and so there must be an isomorphic path in the other model, whose end-point is an appropriate match for P.

We need to fulfil one more requirement of our invariant. Since both models are infinitely multiplied (*we use this feature only here*), and we have only matched up finite subtrees so far, the preceding path can be chosen so that P keeps a distance $> 2^{N-i-1}$ from all nodes that are already matched.

In summary, after n rounds, this matching procedure always produces a partial isomorphism and thus, it is a winning strategy for the Duplicator. ■

The modal invariance theorem can now be proved by chasing a different diagram:

Proof. (Second Proof for the Modal Invariance Theorem)

$$\begin{tikzcd}
\mathfrak{M}, w \arrow[r, "\equiv"'] & \mathfrak{N}, v \arrow[d, "\underline{\leftrightarrow}"] \\
\mathfrak{M}^*, w \arrow[u, "\underline{\leftrightarrow}"] & N^*, v \arrow[l, "\equiv_{\text{FOL}}"]
\end{tikzcd}$$

Again, we arrive at two models $\mathfrak{M}, w$ and N, v that are modally equivalent, and $\mathfrak{N}, v \vDash \varphi(x)$. By lemma 3.7, there exist $\mathfrak{M}^*, w$ and $\mathfrak{N}^*, v$, such that: (i) $\mathfrak{M}, w$ is bisimilar to $\mathfrak{M}^*, w$; (ii) $\mathfrak{N}, v$ is bisimilar to $\mathfrak{N}^*, v$; and (iii) $\mathfrak{M}^*, w$ and $\mathfrak{N}^*, v$ are first-order equivalent. The diagram chasing is now the following: Again we start with $\mathfrak{N}, v \vDash \varphi(x)$. Since $\varphi(x)$ is invariant for bisimulation, $\mathfrak{N}^*, v \vDash \varphi(x)$. Then $\mathfrak{M}^*, w \vDash \varphi(x)$ since $\varphi(x)$ is a first order formula and preserved under first order equivalence. Finally by invariance for bisimulation, $\mathfrak{M}, w \vDash \varphi(x)$.

■

4 Modal Lindström Theorem

An important result in first order model theory is Lindström's characterization of first-order logic. It states that, given a suitable explication of what 'abstract logic' is, first-order logic is the strongest abstract logic to possess the compactness and Löwenheim-Skolem properties:

Theorem 4.1. ***(Lindström's Theorem for FO)****. For any abstract logic L, if $FO \subseteq L$ and L satisfies the Compactness Theorem and the Löwenheim-Skolem Theorem, then $FO = L$.*

Can we have a similar result for modal logic? For instance, can we say that, for any abstract logic L, if $MO \subseteq L$ and L satisfies the Compactness Theorem and the Löwenheim-Skolem Theorem, then $MO = L$? This clearly is not true: FO is a counterexample to this formulation. To see how to formulate Linström theorem for modal logic, consider an alternative formulation of Lindström's theorem for first order logic:

Theorem 4.2. ***(Lindström's Theorem for FO)****. For any abstract logic L, if $FO \subseteq L$ and L satisfies the Compactness Theorem and the Karp property, then $FO = L$.*

The Karp Property says that all formulas of L are invariant for potential isomorphism. A potential isomorphism between two models $\mathfrak{M}$ and $\mathfrak{N}$ is a non-empty family PI of finite partial isomorphisms satisfying two back-and-forth clauses: (a) for any partial isomorphism $F \in PI$ and any d in the domain of $\mathfrak{M}$, there exists e in the domain of $\mathfrak{N}$ such that $F \cup \{(d, e)\} \in PI$, (b) analogously in the opposite direction.

This formulation of Lindström's Theorem for FO suggests a way of formulating Linström's Theorem for ML: We just replace the karp property with a 'karp-like' property, namely invariance for bisimulation. To prove Lindström's Theorem for the basic modal logic ML, we introduce a few definitions and lemmas:

Definition 4.3. ***(Relativization)****. A logic $\mathcal{L}$ has relativization if, for any $\mathcal{L}-$formula φ and new unary proposition letter p (that is, p is irrelevant to the truth value of φ), there is an $\mathcal{L}-$formula $(\varphi)^p$ such that $\mathfrak{M}, w \vDash (\varphi)^p$ iff $\mathfrak{M}|p, w \vDash \varphi$: $\mathfrak{M}|p$ is the submodel of $\mathfrak{M}$ with just the point in $\mathfrak{M}$ where p is true.*

In the proof of Linström's Theorem for ML, we assume that any abstract modal logic has relativization. This assumption can be varied, but it has many natural motivations.

Definition 4.4. ***(Finite Depth Property).*** *For any formula φ, there is a natural number k such that, for all models, $\mathfrak{M}, w \vDash \varphi$ iff $\mathfrak{M}|k, w \vDash \varphi$, where $\mathfrak{M}|k$ is the submodel of $\mathfrak{M}$ with its domain restricted to points reachable from w in k or fewer steps.*

Definition 4.5. ***(n-Bisimulation).*** *Two pointed models $\mathfrak{M}, w$ and $\mathfrak{N}, v$ are $n-$bisimilar (notation: $\mathfrak{M}, w \underline{\leftrightarrow}_n \mathfrak{N}, v$) iff there exists a sequence of binary relations $Z_n \subseteq \cdots \subseteq Z_0$ with the following properties ($i + 1 \leq n$):*

1. *wZ_nv.*
2. *If xZ_0y then x and y agree on all the proposition letters.*
3. *If $xZ_{i+1}y$ and $R^{\mathfrak{M}}xx'$, then there exists y' with $R^{\mathfrak{N}}yy'$ and $x'Z_iy'$.*
4. *If $xZ_{i+1}y$ and $R^{\mathfrak{N}}yy'$, then there exists x' with $R^{\mathfrak{M}}xx'$ and $x'Z_iy'$.*

Lemma 4.6. *If an abstract modal logic $\mathcal{L}$ extends ML, is compact and invariant for bisimulation, then $\mathcal{L}$ has the Finite Depth Property.*

Proof. Let φ be any formula in L, and let p be a new proposition letter that is irrelevant to the truth value of φ. Since we assume that $ML \subseteq L$, $\{\Box^n p \mid n \text{ is a natural number}\} \subseteq L$. We first show that

$$\{\Box^n p \mid n \text{ is a natural number}\} \vDash \varphi \leftrightarrow (\varphi)^p$$

Suppose $\mathfrak{M}, w \vDash \{\Box^n p \mid n \text{ is a natural number}\}$. We focus on the generated sub-model $\mathfrak{M}_w, w$ consisting of w and all the points finitely reachable from it. Clearly, the identity relation is a bisimulation between any pointed model and its generated sub-model. Hence we have that $\mathfrak{M}, w \underline{\leftrightarrow} \mathfrak{M}_w, w$ and that $\mathfrak{M}, w \equiv \mathfrak{M}_w, w$. On the other hand, since p is true in w and all the points finitely reachable from it, p is true in the whole generated sub-model $\mathfrak{M}_w, w$. Therefore, it is easy to see that $\mathfrak{M}_w, w$ is also a generated sub-model of $\mathfrak{M}|p, w$, and so we have that $\mathfrak{M}_w, w \underline{\leftrightarrow} \mathfrak{M}|p, w$ and that $\mathfrak{M}_w, w \equiv \mathfrak{M}|p, w$. Hence $\mathfrak{M}, w \equiv \mathfrak{M}|p, w$, and it follows that

$$\begin{aligned} \mathfrak{M}, w \vDash \varphi \;\; &\text{iff} \;\; \mathfrak{M}|p, w \vDash \varphi && \text{(above)} \\ &\text{iff} \;\; \mathfrak{M}, w \vDash (\varphi)^p && \text{(def. of relativization)} \end{aligned}$$

That is, $\mathfrak{M}, w \vDash \varphi \leftrightarrow (\varphi)^p$.

Next, by applying compactness, we know that there exists a number k such that

$$\{\Box^n p \mid n \leq k\} \vDash \varphi \leftrightarrow (\varphi)^p$$

Let $\mathfrak{M}, w$ be an arbitrary pointed model, and $\mathfrak{M}^*, w$ be the same model except that $V(p)$ is all the points reachable from w in k or fewer steps. Clearly $\mathfrak{M}^*, w \vDash \{\Box^n p \mid n \leq k\}$, and so $\mathfrak{M}^*, w \vDash \varphi$ iff $\mathfrak{M}^*, w \vDash (\varphi)^p$ iff $\mathfrak{M}^*|k, w \vDash \varphi$. Since p is assumed to be new (irrelevant to the truth of φ), we have $\mathfrak{M}, w \vDash \varphi$ iff $\mathfrak{M}|k, w \vDash \varphi$, which is the finite depth property. ■

Lemma 4.7. *If a $L-$formula φ has the Finite Depth Property for distance k, then φ is k-bisimulation invariant.*

Proof. (We merely provide a sketch.) Let two models $\mathfrak{M}, w$ and $\mathfrak{N}, v$ be k-bisimilar, and $\mathfrak{M}, w \vDash \varphi$ for a $L-$formula φ that has the finite depth property. By the standard unraveling technique, the two models are bisimilar to their tree unraveling: $\mathfrak{M}, w \underline{\leftrightarrow}$ *Tree*$(\mathfrak{M}), w^*$ and $\mathfrak{N}, v \underline{\leftrightarrow}$ *Tree*$(\mathfrak{N}), v^*$. Now, we can show that (1) *Tree*$(\mathfrak{M}), w^*$ and *Tree*$(\mathfrak{N}), v^*$ are $k-$bisimilar, and therefore (2) *Tree*$(\mathfrak{M})|k, w^*$ and *Tree*$(\mathfrak{N})|k, v^*$ are bisimilar. Hence,

$$\begin{aligned}
\mathfrak{M}, w \vDash \varphi \ \text{ iff } \ & \text{Tree}(\mathfrak{M}), w^* \vDash \varphi && \text{(invariance)} \\
\text{iff } \ & \text{Tree}(\mathfrak{M})|k, w^* \vDash \varphi && \text{(finite depth property)} \\
\text{iff } \ & \text{Tree}(\mathfrak{N})|k, v^* \vDash \varphi && \text{(invariance)} \\
\text{iff } \ & \text{Tree}(\mathfrak{N}), v^* \vDash \varphi && \text{(finite depth property)} \\
\text{iff } \ & \mathfrak{N}, v \vDash \varphi && \text{(invariance)}
\end{aligned}$$

■

Lemma 4.8. *If an $L-$formula φ is k-bisimulation invariant, then φ is definable by a modal formula of modal operator depth k.*

Proof. (Again, we just provide a sketch.) Let φ be a $L-$formula that is $k-$bisimilar invariant. We know that in the basic modal logic, there are only finitely many non-equivalent basic modal formulas of degree at most k. Let Γ_k be the set of all these basic modal formulas. To show that φ is definable by a modal formula of degree at most k, it suffices to show the following:

If $\mathfrak{M}, w$ and $\mathfrak{N}, v$ agree on all formulas in Γ_k, then they agree on φ.

For then, φ will be equivalent to a boolean combination of formulas in Γ_k. To show this, it suffices show the following fact:

> **Fact**: If $\mathfrak{M}, w$ and $\mathfrak{N}, v$ agree on all formulas of degree at most k, then $\mathfrak{M}, w$ and $\mathfrak{N}, v$ are $k-$bisimilar.

We prove this fact by defining a sequence of binary relations $Z_k \subseteq \cdots \subseteq Z_0$ as follows: for all $0 \leq i \leq k$, xZ_iy iff x and y agree on all modal formulas of degree at most i. We show that this sequence of binary relations is a n-bisimulation between $\mathfrak{M}, w$ and $\mathfrak{N}, v$.

The first two conditions of the n-bisimulation follow immediately from the above definition of the sequence. For the forth condition, suppose $xZ_{i+1}y$, that is, x and y agree on all modal formulas of degree at most $i+1$. Also, suppose that $R^{\mathfrak{M}}xx'$, and let Γ_i be the set of all modal formulas of degree i that are true at x'. Since Γ_i is a finite set, $\bigwedge \Gamma_i$ is a modal formula. Then $\Diamond \bigwedge \Gamma_i$ is of degree at most $i+1$, and $\mathfrak{M}, x \vDash \Diamond \bigwedge \Gamma_i$. Since x and y agree on all modal formulas of degree at most $i+1$, $\mathfrak{N}, y \vDash \Diamond \bigwedge \Gamma_i$, and so there exists y' with $R^{\mathfrak{N}}yy'$ such that $\mathfrak{N}, y' \vDash \bigwedge \Gamma_i$. But then x' and y' agree on all modal formulas of degree at most i, that is, we have $x'Z_iy'$. The proof for the back condition is similar. ■

Theorem 4.9. ***(Modal Lindström Theorem).*** *If an abstract modal logic L extends ML, is compact and invariant for bisimulation, then $L = ML$.*

Proof. An immediate consequence of Lemma 4.6, 4.7 and 4.8 ■

It is sometimes complained that Lindström's Theorem has no concrete applications. As a rebuttal, we now derive the modal invariance theorem from the modal Lindström's theorem:

Theorem 4.10. *The Modal Lindström Theorem implies the Modal Invariance Theorem.*

Proof. Suppose $\varphi(x)$ is a first order formula that is invariant for bisimulation. Define an abstract logic L by adding φ to the basic modal language, and then closing off the result (in some suitable syntax) under (a) Boolean operations, (b) existential modalities $\Diamond$, and (c) an operation of relativiation α^β where α, β are already formulas in the language. This language contains the basic modal language and can also be translated into a fragment of first order logic, and so it is compact (due to the compactness of first-order logic).

Next, we prove that all $L-$formulas are invariant for bisimulation by induction on $L-$formulas. Since formulas of modal logic are bisimulation invariant and φ is bisimulation invariant by assumption, the only inductive case that needs checking is the relativization formula α^β, where we already assume bisimulation invariance for α, β. Suppose E is a bisimulation between two models $\mathfrak{M}, w$ and $\mathfrak{N}, v$, while $\mathfrak{M}, w \vDash \alpha^\beta$. By definition, we have $\mathfrak{M}|\beta, w \vDash \alpha$. We observe that the relation $E|\beta$ consists of all pairs in E which connect $\beta-$worlds in $\mathfrak{M}$ to $\beta-$worlds in $\mathfrak{N}$ is itself a bisimulation between $\mathfrak{M}|\beta, w$ and $\mathfrak{N}|\beta, v$: To check the zigzag clause, suppose that $\mathfrak{M}, x \vDash \beta$, $\mathfrak{N}, y \vDash \beta$ and xEy. Let $R^{\mathfrak{M}}uu'$ in $\mathfrak{M}$ with $\mathfrak{M}, u' \vDash \beta$. Since E is a bisimulation, there exists a world v' in $\mathfrak{N}$ with $R^{\mathfrak{N}}vv'$ and $u'Ev'$. But β is assumed to be invariant for bisimulation; hence $\mathfrak{N}, v' \vDash \beta$. Thus, we have shown the zigzag property for the relation $E|\beta$, and so $E|\beta$ is a bisimulation between $\mathfrak{M}|\beta, w$ and $\mathfrak{N}|\beta, v$. Since α is invariant for bisimulation and $\mathfrak{M}|\beta, w \vDash \alpha$, we have $\mathfrak{N}|\beta, v \vDash \alpha$; that is, $\mathfrak{N}, \beta \vDash \alpha^\beta$.

Summarizing, we have shown that the abstract logic L satisfies all the conditions of Modal Lindström's Theorem, and hence $L = ML$. In particular, φ is equivalent to the standard translation of a modal formulas. ■

Both the modal invariance theorem and the modal Lindström theorem suggest that modal logic is a special fragment of first order logic. However, consider the following theorem in temporal logic:

Theorem 4.11. ***(Kamp's Theorem).*** *On complete linear orders, the full first order logic of $\{<, \vec{P}\}$ is equivalent (in express power) with the temporal logic of $\{Since, Until\}$.*

Does this theorem contradict what we said earlier, namely that first order logic is more expressive than modal logic/temporal logic? Not so. This is because Kamp's result only holds for a particular class of models—complete linear orders—rather than arbitrary models. On special models, it may be the case that first order logic has the same expressive power as temporal logic (there are even more cases for this than complete linear orders), but it is not the case for arbitrary models.

A related observation behind this expressive completeness is the following: on complete linear orders, the full first order logic is equivalent to the full first order logic with only *three variables*, free or bound. (References for these results can be found in the literature on temporal logic, a brief introduction is found in van Benthem [21].)

5 Sahlqvist Theorem

So far we have focused on a few important results about modal logic from the perspective of models: For instance, modal invariance theorem says that modal formulas cannot tell the difference between bisimilar models, and that any first order formula which also cannot make such distinctions is a modal formula. The modal Lindström Theorem suggests that compactness and invariance under bisimulations in some sense characterize modal logic.

However, we can also understand modal formulas as asserting something about the underlying frame. On such a frame-based perspective, we consider special axioms and classes of frames on which they are valid. For instance, $\Box\varphi \to \Box\Box\varphi$ is valid on a frame $\mathfrak{F}$ if and only if $\mathfrak{F}$ is transitive. Can we say something systematic about the relation between the syntactic shape of the axioms and the frame properties they correspond to? In general, the frame truth a modal formula is equivalent to a second order formula. In reality, however, many of these second order formulas are also equivalent to first order formulas. ($\Box\varphi \to \Box\Box\varphi$ is just one example.) Fitch, for instance, observed that all model axioms of the form $\Diamond^k\Box^m\varphi \to \Box^i\Diamond^j\varphi$ have first order correspondents.

A more general result is due to Sahlqvist (we do not go into the precise history here):

Theorem 5.1. ***(Sahlqvist Theorem).*** *There is an effective method for computing the first order correspondent of any formulas of the form $\alpha \to \beta$:*

$\alpha :\ p \mid \Box^k p\ (k \in \mathbb{N}) \mid \Diamond \mid\ \wedge\ \mid \vee$

$\beta :\ p \mid \Diamond \mid \Box \mid\ \wedge\ \mid \vee$

Proof. (We provide a sketch containing the main ideas.) By distributing $\Diamond$ over $\vee$, we can turn α into an equivalent disjunction, each of its disjunct is constructed out of $\Diamond \mid \wedge \mid \Box^k p$. Since α is the antecedent of a conditional, we can get rid of the disjunction by means of the following equivalence:

$$(\varphi \vee \psi) \to \gamma \Longleftrightarrow (\varphi \to \gamma) \wedge (\psi \to \gamma)$$

and so we only need to consider α that are constructed out of $\Diamond| \wedge |\Box^k p$. Now, consider the second-order translation $\forall \vec{P}(ST(\alpha) \to ST(\beta))$:

Step 1. If there are diamonds in α, pull out the corresponding existential quantifiers in $ST(\alpha)$, using equivalences of the form:

$$(\exists x \varphi(x) \wedge \psi) \Longleftrightarrow \exists x(\varphi(x) \wedge \psi),$$

$$(\exists x \varphi(x) \to \psi) \Longleftrightarrow \forall x(\varphi(x) \to \psi),$$

Step 1 results in a formula of the form $\forall \vec{P} \forall \vec{y}$(conjunctions of translations of $\Box^k p \to ST(\beta)$).

Step 2 (Miminal valuation). For formulas of the form $\Box^k p$, there are minimal valuations that make them true. For instance, in order to make $\Box p$ true at w, it is sufficient to make p true at all the accessible worlds from w. For each proposition letter p, its minimal valuation is a first order formulas α_p.

Step 3 (Instantiation) .Think of an implication as a promise: If you give me the minimal way of making the antecedent true, you will get the consequent. By substituting the minimal valuations for each of the predicate P in the consequent $ST(\beta)$, we arrive at a first order formula

$$\forall \vec{y}[\overline{\alpha_p/P}]ST(\beta)$$

which is the first order correspondent of $\alpha \to \beta$.

Next, we show that the algorithm works, that is, $\mathfrak{F} \vDash \alpha \to \beta$ iff $\mathfrak{F} \vDash \forall \vec{y}[\overline{\alpha_p/P}]ST(\beta)$.

($\to$): This is the easy direction. If $\mathfrak{F} \vDash \alpha \to \beta$, then $\alpha \to \beta$ holds for all the valuations. In particular, it holds for the minimal valuation. In short, the $\to$ direction is just universal instantiation in second order logic:

$$\forall P \varphi(P) \to \varphi(\alpha_p/P)$$

($\leftarrow$): Let V be an arbitrary valuation and suppose $(\mathfrak{F}, V), w \vDash \alpha$, we want to show that

$(\mathfrak{F}, V), w \vDash \beta$. If $(\mathfrak{F}, V), w \vDash \alpha$, then there exist some valuations that make α true, and so there exists a minimal valuation $V_{\min}$ that makes α true: $(\mathfrak{F}, V_{\min}), w \vDash \alpha$. Hence, the minimal valuation also makes β true: $(\mathfrak{F}, V_{\min}), w \vDash \forall\vec{y}[\overline{\alpha_p/P}]ST(\beta)$, which is equivalent to $(\mathfrak{F}, V_{\min}), w \vDash \beta$. $V_{\min}$ is not necessarily the same as V, because V may assign larger sets to the predicates. However, given our assumption that β is a positive formula, it retains its truth value when the valuations of its predicates are made larger. Therefore, $(\mathfrak{F}, V), w \vDash \beta$. ■

What happens if we restrict to particular classes of frames (such as transitive frames)? Then more modal formulas will become first order definable. As we will see, the McKinsey Axiom

$$\Box\Diamond p \to \Diamond\Box p$$

is not first order definable. On transitive frames, however, the McKinsey axiom is indeed first order definable, and it expresses atomicity of the ordering: every point has a reflexive endpoint above it. (Alternatively: $(\Box p \to \Box\Box p) \wedge (\Box\Diamond p \to \Diamond\Box p)$ is first order definable on arbitrary frames.) The proof of this result involves the Axiom of Choice in an essential way, making it different from the above Sahlqvist-style minimal valuation reasoning. As a background to this observation, a general result shows that all modal reduction principles are first-order definable on transitive frames.

6 Modal formulas without FO Correspondents

Let us consider how to prove that certain modal formulas are not first order definable. We consider two examples: Löb's formula $\Box(\Box p \to p) \to \Box p$, and McKinsey's formula $\Box\Diamond p \to \Diamond\Box p$.

6.1 Löb's formula

Theorem 6.1. *Löb's formula $\Box(\Box p \to p) \to \Box p$ does not correspond to a first order condition.*

Proof. We first show that Löb's formula defines the class of frames (W, R) such that R is transitive and R's converse is well-founded (that is, there is no infinite ascending chain $x_0 R x_1 R x_2 R ...$).

Suppose $\mathfrak{F} = (W, R)$ is a frame with a transitive and conversely well-founded relation R, and then suppose for contradiction that Löb's formula is not valid in $\mathfrak{F}$. This means that there is a valuation V and a state w such that $(\mathfrak{F}, V), w \nvDash \Box(\Box p \to p) \to \Box p$. That is, $(\mathfrak{F}, V), w \vDash \Box(\Box p \to p)$ but $(\mathfrak{F}, V), w \nvDash \Box p$. Then w must have a successor w_1 such that $w_1 \nvDash p$, and as $\Box p \to p$ holds at all successors of w, we have $w_1 \not\models \Box p$. This in turn implies that w_1 has a successor w_2 where p is false; by the transitivity of R, w_2 is a successor of w. By repeating this argument, we can find an infinite path $wRw_1Rw_2R\ldots$, contradicting the converse well-foundedness of R.

For the other direction, suppose that either R is not transitive or its converse is not well-founded; in both cases we have to find a valuation R and a state w such that $(\mathfrak{F}, V), w \not\models \Box(\Box p \to p) \to \Box p$. We focus on the case where R is transitive but not conversely well-founded. That is, suppose we have a transitive frame containing an infinite sequence $w_0Rw_1Rw_2R\ldots$. We then define a valuation v as follows:

$$V(p) = W - \{x \in W \mid \text{there is an infinite path starting from } x\}$$

It is straightforward to verify that $(\mathfrak{F}, V), w \models \Box(\Box p \to p)$ and $(\mathfrak{F}, V), w \not\models \Box p$.

Next, we show that the class of transitive and conversely well-founded frames cannot be defined in the first order language by using a compactness argument. Suppose for contradiction that there is a first order formula φ equivalent to Löb's formula. Hence any model making φ true must be transitive. Let $\sigma_n(x_0, \ldots, x_n)$ be the first order formula stating that there is an $R-$path of length n through $x_0, \ldots, x_n$:

$$\sigma_n(x_0, \ldots, x_n) = \bigwedge_{0 \le i < n} Rx_ix_{i+1}$$

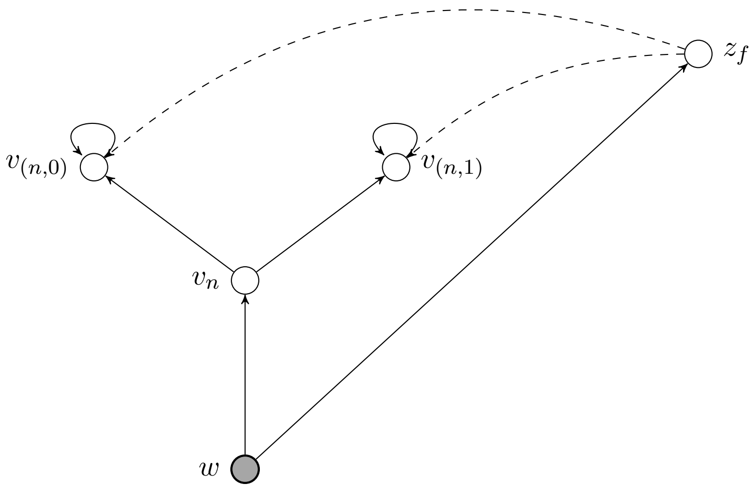
Then, every finite subset of

$$\Sigma = \{\varphi\} \cup \{\forall xyz((Rxy \wedge Ryz) \to Rxz)\} \cup \{\sigma_n \mid n \in \omega\}$$

is satisfiable in a finite linear order, and hence in the class of transitive, conversely well-founded frames. Thus by the compactness theorem for first order logic, Σ itself has a model. But any model of Σ is not conversely well-founded, contradicting the assumption that φ defines the class of transitive, conversely well-founded frames. Hence Löb's formula cannot be equivalent to any first order formula. ■

6.2 McKinsey formula

Theorem 6.2. *McKinsey formula $\Box\Diamond p \to \Diamond\Box p$ does not correspond to a first-order condition.*



Proof.

Consider the frame $\mathfrak{F} = (W, R)$ such that

$$W = \{w\} \cup \{v_n, v_{(n,i)} \mid n \in \mathbb{N}, i \in \{0,1\}\} \cup \{z_f \mid f : \mathbb{N} \to \{0,1\}\}$$

$$R = \{(w, v_n), (v_n, v_{(n,i)}), (v_{(n,i)}, v_{(n,i)}) \mid n \in \mathbb{N}, i \in \{0,1\}\} \cup \{(w, z_f), (z_f, v_{(n,f(n))}) \mid n \in \mathbb{N}, f : \mathbb{N} \to \{0,1\}\}$$

We first show that $\mathfrak{F}, w \vDash \Box\Diamond p \to \Diamond\Box p$. Suppose that $(\mathfrak{F}, V), w \vDash \Box\Diamond p$. What this means is that each v_n has a p-successor, that is, either $(\mathfrak{F}, V), v_{(n,0)} \vDash p$ or $(\mathfrak{F}, V), v_{(n,1)} \vDash p$. But then there exists a choice function $f : \mathbb{N} \to \{0,1\}$ such that $(\mathfrak{F}, V), v_{(n,f(n))} \vDash p$ for every $n \in \mathbb{N}$. Then z_f is a witness of $\Box p$, and so w is a witness of $\Diamond\Box p$—that is, $(\mathfrak{F}, V), w \vDash \Diamond\Box p$.

Next, we show that $\Box\Diamond p \to \Diamond\Box p$ is not first order definable. By the downward Löwenheim-Skolem Theorem, there exist a countable elementary submodel $\mathfrak{F}^-$ of $\mathfrak{F}$ whose domain W^- contains w, and each v_n, $v_{(n,0)}$ and $v_{(n,1)}$. As W is uncountable and W^- is countable, there must be a choice function $f : \mathbb{N} \to \{0,1\}$ such that $z_f \notin W^-$. Now, if the McKinsey formula was equivalent to a first formula φ, then since $\mathfrak{F}, w \vDash \varphi$, it follows that $\mathfrak{F}^-, w \vDash \varphi$ and that $\mathfrak{F}^-, w \vDash \Box\Diamond p \to \Diamond\Box p$. But we will show that $\mathfrak{F}^-, w \nvDash \Box\Diamond p \to \Diamond\Box p$, hence the McKinsey formula cannot be equivalent to a first order formula.

Let f be a choice function such that $z_f \notin W^-$, and define V^- as a valuation on F^- such that $V^-(p) = \{v_{(n,f(n))} \mid n \in \mathbb{N}\}$. We will show that $(\mathfrak{F}^-, V^-), w \vDash \Box\Diamond p$ and that $(\mathfrak{F}^-, V^-), w \nvDash \Diamond\Box p$.

To see that $\Box\Diamond p$ is true at w, note first that p is true at exactly one of $v_{(n,0)}$ and $v_{(n,1)}$ for every n, and so $\Diamond p$ is true at v_n for every n. We still need to show that $\Diamond p$ is true at an arbitrary

$z_g \in W^-$, and it suffices to show that there exists some $n \in \mathbb{N}$ such that $g(n) = f(n)$. Suppose otherwise; then $g(n) = 1 - f(n)$. But such a relation is expressible in first order logic and preserved under elementary submodels, and so it follows that if $z_g \in W^-$, then $z_f \in W^-$, contradicting our earlier assumption that $z_f \notin W^-$. Hence there exists some $n \in \mathbb{N}$ such that $g(n) = f(n)$; that is, z_g has a p successor $v_{(n,g(n))}$.

Finally, it is easy to see that $\Diamond\Box p$ is false at w. For a start, since p is true at exactly one of the successor of v_n, $\Box p$ is false at v_n for every n. Moreover, for any z_g in W^-, $\Box p$ is false at z_g as well: Since g is different from f, there exists at least one $n \in \mathbb{N}$ such that $g(n) \neq f(n)$, and so z_g has at least one successor $v_{(n,g(n))}$ where p is false. ■

Now we have seen modal formulas that are first order definable (Sahlqvist) and modal formulas that are not (Löb, McKinsey). Actually there are differences among formulas that are not first order definable: Löb's formula, for instance, is much better behaved than McKinsey's formula. What this means will become clear when we discuss fixed-point logics.

7 More on First-Order Correspondence

In the first half of the notes, we discussed two perspectives on modal logic: A modal-based perspective is provided by the standard translation. We discussed two major theorems from this perspective: The Invariance Theorem and Linström's Theorem. A second perspective is the frame-based perspective. We discussed Sahlqvist Theorem and the first order undefinability of Löb's axiom and McKinsey formula.

We can add a more general question to the discussion of Sahlqvist Theorem: Given a modal formula, can we determine whether its frame truth is first order? The answer to this question is: It is at least undecidable. What this means is that fairly complicated questions can be encoded as correspondence questions. We will also give the following proof:

Theorem 7.1. *First order definability of monadic Π^1_1 sentences is not arithmetical.*

Proof. Let φ be a monadic Π^1_1 formula that defines $(V_\omega, \in)$ up to isomorphism, and let α be a first order formula in the language of set theory. We can show that:

$$\varphi \models \alpha \text{ if and only if } \varphi \vee \alpha \text{ is first order definable.}$$

If we can show this, then we are done, because we can then reduce the left hand side to the right hand side, and the problem on the left hand side is not arithmetical: $\varphi \models \alpha$ if and only if α is true in $(V_\omega, \in)$, and truth in $(V_\omega, \in)$ has the same complexity as truth in arithmetic, which is not arithmetical by Tarski's Undefinability Theorem.

($\Longrightarrow$): If $\varphi \models \alpha$, then $\varphi \vee \alpha$ is equivalent to the first order formula α.

($\Longleftarrow$): Suppose $\varphi \vee \alpha$ is equivalent to a first order formula β. Assume $(V_\omega, \in) \models \neg\alpha$. We know that $(V_\omega, \in) \models \varphi$, and so $(V_\omega, \in) \models \varphi \vee \alpha$, hence $(V_\omega, \in) \models \beta$. By the upward Skolem-Lowenheim Theorem, there is some uncountable elementary extension $(V^+, \in)$ of $(V_\omega, \in)$. We have $(V^+, \in) \models \beta$ because β is a first order formula. Hence $(V^+, \in) \models \varphi \vee \alpha$. But $(V^+, \in) \not\models \alpha$, since $(V_\omega, \in) \not\models \alpha$ and α is a first order formula. Therefore, $(V^+, \in) \models \varphi$. But that cannot be, because φ is supposed to define $(V_\omega, \in)$ up to isomorphism. ∎

What this result suggests is that definability questions can be difficult in second order logic. Since modal logic is a special fragment of monadic Π^1_1, the above result does not immediately apply, but Lydia Chagrova has shown that, in fact, first-orderness of modal axioms is undecidable.

8 More Background in Model Theory

There is a famous result in first-order model theory that characterizes elementary classes (classes of models that are first-order definable) in terms of the closure properties the classes needs to have:

Definition 8.1. *A class of models* K *is* ***elementary*** *if* $\mathsf{K} = \mathrm{Mod}(\Sigma)$ *for some set of first order formulas* Σ*, where* $\mathrm{Mod}(\Sigma)$ *is the class of all models of* Σ*.*

Theorem 8.2. K *is an elementary class if and only if* K *is closed under isomorphisms, ultraproducts, and* $\overline{\mathsf{K}}$ *(the complement of* K*) is closed under ultrapowers.*

In this section, we will prove a similar result concerning modal logic, with 'elementary class' replaced by 'modally definable class' and with 'isomorphism' replaced by 'bisimulation'. Our result will help answering the question: Which properties of *models* are definable by means of modal formulas? First, we prove a useful lemma:

Lemma 8.3. *Let* Σ *be a set of modal formulas, and* K *a class of pointed models in which* Σ *is finitely satisfiable. Then* Σ *is satisfiable in some ultraproduct of models in* K*.*

Proof. Define an index set I as the collection of all finite subsets of Σ:

$$I = \{\Sigma_0 \subseteq \Sigma \mid \Sigma_0 \text{ is finite}\}$$

We construct an ultrafilter U over I as follows: For each $\sigma \in \Sigma$, let $\hat{\sigma}$ be the set of all $i \in I$ such that $\sigma \in i$. Then the set $E = \{\hat{\sigma} \mid \sigma \in \Sigma\}$ has the finite intersection property because

$$\{\sigma_1, \ldots, \sigma_n\} \in \hat{\sigma_1} \cap \cdots \cap \hat{\sigma_n}$$

By corollary 2.4, E can be extended to an ultrafilter U over I.

Moreover, Σ is finitely satisfiable in K, which means that for each $i \in I$, there exists a pointed model $\mathfrak{N}_i, v_i$ in K such that $\mathfrak{N}_i, v_i \vDash i$. Hence we can define the ultraproduct $\prod_U \mathfrak{N}_i$. We also define a state f_U as follows: Let W_i be the universe of the model $\mathfrak{N}_i$ and consider the function $f \in \prod_{i \in I} W_i$ such that $f(i) = v_i$. We now show that $\prod_U \mathfrak{N}_i, [f]_U \vDash \Sigma$.

It is easy to see that if $i \in \hat{\sigma}$, then $\sigma \in i$ and so $\mathfrak{N}_i, v_i \vDash \sigma$. Hence for each $\sigma \in \Sigma$,

$$\{i \in I \mid \mathfrak{N}_i, v_i \vDash \sigma\} \supseteq \hat{\sigma}$$

Since $\hat{\sigma} \in U$ and U is an ultrafilter, $\{i \in I \mid \mathfrak{N}_i, v_i \vDash \sigma\} \in U$. By the fundamental theorem of ultraproducts, $\prod_U \mathfrak{N}_i, [f]_U \vDash \Sigma$. ■

Definition 8.4. *A class K of pointed models is **modally definable** if $\mathsf{K} = \mathrm{Mod}(\Sigma)$ for some set of modal formulas Σ; that is, for any pointed model $(\mathfrak{M}, w)$, we have $(\mathfrak{M}, w) \in \mathsf{K}$ iff for all $\sigma \in \Sigma$, $\mathfrak{M}, w \vDash \sigma$.*

Theorem 8.5. *A class K of pointed models is modally definable iff K is closed under bisimulations, ultraproducts, and $\bar{\mathsf{K}}$ is closed under ultrapowers.*

Proof. ($\Longleftarrow$) Let T be the modal theory of K:

$$T = \{\varphi \mid \text{for all } (\mathfrak{M}, w) \text{ in } \mathsf{K}, (\mathfrak{M}, w) \vDash \varphi\}$$

Clearly, we have that $\mathsf{K} \subseteq \mathrm{Mod}\, T$. We now show that, in fact, $\mathrm{Mod}\, T \subseteq \mathsf{K}$.

Suppose that $\mathfrak{M}, w \vDash T$. Our goal is to show that $\mathfrak{M} \in \mathsf{K}$ using the closure properties.

Let Σ to be the modal theory of w; that is, $\Sigma = \{\varphi \mid \mathfrak{M}, w \vDash \varphi\}$. Suppose Δ is a finite subset of Σ. Then there must be a model $\mathsf{K}_\Delta \in \mathsf{K}$ which satisfies Δ: Otherwise, $\neg \bigwedge \Delta \in T$, and so $\mathfrak{M}, w \vDash \neg \bigwedge \Delta$, contradicting $\mathfrak{M}, w \vDash \Delta$. Hence Σ is finitely satisfiable in K. It follows from lemma 8.3 and the closure of K under taking ultraproducts that Σ is satisfiable in some model $\mathfrak{N}, v$ in K. But $N, v \vDash \Sigma$ implies that $\mathfrak{N}, v$ and $\mathfrak{M}, w$ are modally equivalent. If we take the ultrapowers of $\mathfrak{M}, w$ and $\mathfrak{N}, v$ over a countably incomplete ultrafilter U, we get $\prod_U \mathfrak{M}, w^+$ and $\prod_U \mathfrak{N}, v^+$, both of which are saturated and are elementary extensions of $\mathfrak{M}, w$ and $\mathfrak{N}, v$ respectively by corollary 2.15. Moreover, by lemma 3.2, we know that $\prod_U \mathfrak{M}, w^+ \underline{\leftrightarrow} \prod_U \mathfrak{N}, v^+$. The following diagram illustrates the whole process of model constructions:

$$\begin{tikzcd}
\mathfrak{M}, w \arrow[r, "\equiv"'] & \mathfrak{N}, v \arrow[d] \\
\prod_U \mathfrak{M}, w^+ \arrow[u] & \prod_U \mathfrak{N}, v^+ \arrow[l, "\underline{\leftrightarrow}"]
\end{tikzcd}$$

Now we start diagram chasing: We know that $(\mathfrak{N}, v) \in \mathsf{K}$. K is closed under ultrapowers, hence $(\prod_U \mathfrak{N}, v^+) \in \mathsf{K}$. Since $(\prod_U \mathfrak{M}, w^+)$ and $(\prod_U \mathfrak{N}, v^+)$ are bisimilar and K is closed under bisimulations, $(\prod_U \mathfrak{M}, w^+) \in \mathsf{K}$. Then $(\mathfrak{M}, w)$ has to be in K as well; otherwise, $(\mathfrak{M}, w) \in \overline{\mathsf{K}}$, and since $\overline{\mathsf{K}}$ is closed under ultrapowers, then the $(\prod_U \mathfrak{M}, w^+)$ is also in $\overline{\mathsf{K}}$, not in K. ■

Comment. One part of the proof here is fairly arbitrary, namely the extension to just some ultrafilter in Lemma 8.3. Intuitively, the initial filter should be sufficient, and the only reason we extend it to an ultrafilter is to be able to use standard model-theoretic results. There are some alternatives for this in so-called ‘possibility semantics’ for modal logics that can work with filters only. These involve a broader bimodal setting with an accessibility relation plus an inclusion relation, but we forego this alternative less classical line here. (For a recent perspective, cf. the cited paper by van Benthem, Bezhanishvili and Holliday, 2015.)

9 Frame-Building Operations

We now switch back to the frame-based perspective and turn out attention to the Goldblatt-Thomason Theorem. This fundamental result characterizes the modal definability of elementary classes of *frames* in terms of closure under four frame-building operations. (Each operation is also an operation that works on models as well as frames, by tacking on the appropriate valuations.) Failure of closure under these frame operations can be useful in showing that certain first order properties are not modally definable.

9.1 Generated Subframes

Definition 9.1. *A frame (W', R') is a* ***generated subframe*** *of (W, R) if and only if*

1. *$W' \subseteq W$,*

2. *For all $x, y \in W'$, $R'xy$ iff Rxy,*

3. *For all $x \in W'$ and $y \in W$, if Rxy then $y \in W'$.*

The model (W', R', V') is a generated submodel of (W, R, V) if in addition $V'(p) = V(p) \cap W'$.

It is straightforward to show that, if $\mathfrak{F}'$ is a generated subframe of $\mathfrak{F}$, then for every modal formula φ, $\mathfrak{F} \vDash \varphi$ implies $\mathfrak{F}' \vDash \varphi$. Hence, going from a frame to a generated subframe preserves the validity of modal formulas. It follows that if a class of frames K is modally definable, then it must be closed under generated subframes: Suppose K is the class of all the frames on which the set Σ of modal formulas is valid. If $\mathfrak{F} \in \mathsf{K}$, Σ is valid on $\mathfrak{F}$, and hence Σ is valid on any generated subframe of $\mathfrak{F}$ as well. So any generated subframe of $\mathfrak{F}$ is also in K.

Therefore, if the validity of some first order sentence is not closed under taking generated subframes, then it doesn't correspond to a modal formula. Take the first order sentence $\exists x Rxx$. We claim that $\mathsf{K} = \{\mathfrak{F} \mid \mathfrak{F} \vDash \exists x Rxx\}$ is not modally definable. For instance, consider a frame $\mathfrak{F}$ with two isolated points x and y, one of which, x, can see itself, while y cannot. Now, $\{y\}$ is a generated subframe of $\mathfrak{F}$, and $\mathfrak{F} \vDash \exists x Rxx$, yet $\{y\} \nvDash \exists x Rxx$.

9.2 Disjoint Unions

Definition 9.2. *For disjoint (that is, sharing no common elements) frames $\mathfrak{F}_i = (W_i, R_i)(i \in i)$, their **disjoint union** is the structure $\biguplus_i \mathfrak{F}_i = (W, R)$, where W is the union of the sets W_i and R is the union of the relations R_i. The disjoint union of models will also preserve the valuation of each model.*

Let $\{\mathfrak{F}_i \mid i \in I\}$ be a family of frames. We can show that if $\mathfrak{F}_i \vDash \varphi$ for every $i \in I$, then $\biguplus_i \mathfrak{F}_i \vDash \varphi$. It follows that if a class of frames K is modally definable, then it must be closed under disjoint unions as well. For instance, the class of finite frames is not modally definable, because it is not closed under disjoint unions: A disjoint union of infinitely many singleton frames is no longer finite. (This class of frames is closed under generated subframes, however.)

9.3 Bounded Morphic Images

Definition 9.3. *A function $f : (W_1, R_1) \to (W_2, R_2)$ between frames is a **bounded morphism** if and only if*

1. *For all $x, y \in W_1$, if R_1xy then $R_2f(x)f(y)$.*
2. *For all $x \in W_1$ and $y \in W_2$, if $R_2f(x)y$, then there exists an $x' \in W_1$ such that R_1xx' and $f(x') = y$.*

*If there is a surjective bounded morphism from $\mathfrak{F}_1$ to $\mathfrak{F}_2$, then we say that $\mathfrak{F}_2$ is a **bounded morphic image** of $\mathfrak{F}_1$. Moreover, f is a bounded morphism between two models (W_1, R_1, V_1) and (W_2, R_2, V_2) if, in addition, $x \in V_1(p)$ if and only if $f(x) \in V_2(p)$ for all $x \in W_1$ and all proposition letter p.*

Taking bounded morphic images also preserves the validity of modal formulas, and so modally definable classes of frames must be closed under taking bounded morphic images. For example, $\mathsf{K} = \{F \mid F \vDash \forall x \neg Rxx\}$ is not modally definable because it fails to be closed under taking bounded morphic images, even though it is closed under generated subframes and disjoint unions: Just take a a frame with two mutually accessible points and a second frame with one reflexive point, and there is a surjective bounded morphism from the first to the second. However, $\forall x \neg Rxx$ is satisfied in the first frame but not in the second.

9.4 Ultrafilter Extensions

Definition 9.4. *Let $\mathfrak{F} = (W, R)$ be a frame. The* ***ultrafilter extension*** $\mathfrak{ue}(\mathfrak{F})$ *of* $\mathfrak{F}$ *is defined as the frame* $(Uf(W), R^{\mathfrak{ue}})$*:*

1. *$Uf(W)$ is the set of ultrafilters over W.*

2. *Let U, V be two ultrafilters over W. $R^{\mathfrak{ue}}UV$ iff $\forall X \in V, m_R(X) \in U$, where $m_R(X) = \{w \in W \mid \exists x \in X : wRx\}$.*

   Alternatively, if we define $l_R(X) = \{w \in W \mid \forall x \in W : \text{ if } wRx \text{ then } x \in X\}$, then $R^{\mathfrak{ue}}UV$ iff $\{X \mid l_R(X) \in U\} \subseteq V$.

The ultrafilter extension of a model $(\mathfrak{F}, V)$ is the model $\mathfrak{ue}(\mathfrak{M}) = (\mathfrak{ue}(\mathfrak{F}), V^{\mathfrak{ue}})$, where $V^{\mathfrak{ue}}(p)$ is the set of ultrafilters of which $V(p)$ is a member.

Several features of the ultrafilter extensions are worth noting. First, the ultrafilter extension is analogous to the canonical frame in the completeness proof, and the ultrafilter extension of a model is analogous to the canonical model: A proposition letter p is true at an ultrafilter U just in case U contains $V(p)$, and We can prove a result analogous to the truth lemma of the completeness proof: For any modal formula φ,

$$\mathfrak{ue}\mathfrak{F}, U \vDash \varphi \text{ iff } V(\varphi) \in U.$$

Second, the ultrafilter extension (of a model) is a saturated model:

Lemma 9.5. *The ultrafilter extension $\mathfrak{ue}\mathfrak{M}$ of a modal model $\mathfrak{M}$ is modally saturated.*

Proof. Let $U \in \mathfrak{ue}\mathfrak{M}$ be an ultrafilter and let Σ be a collection of modal formulas which is finitely satisfiable in the set of successors of U. We would like to find an ultrafilter U' such that $R^{\mathfrak{ue}}UU'$ and $\mathfrak{ue}\mathfrak{M}, U' \vDash \Sigma$. Define

$$\Delta = \{V(\phi) \mid \phi \in \Sigma\} \cup \{X \mid l_R(X) \in U\}$$

We show that Δ has the finite intersection property. Since both $\{V(\phi) \mid \phi \in \Sigma\}$ and $\{X \mid l_R(X) \in U\}$ are closed under taking intersections, it suffices to show that for any $\phi \in \Sigma$ and any $l_R(X) \in U$, we have $V(\phi) \cap X \neq \emptyset$. Since $\phi \in \Sigma$, there must be a successor U'' of U such that $\mathfrak{ue}\mathfrak{M}, U'' \vDash \phi$;

hence $V(\phi) \in U''$. Moreover, $l_R(X) \in U$ implies that $X \in U''$. Hence, $V(\phi) \cap X \in U''$ and so $V(\phi) \cap X \neq \emptyset$, since U'' is an ultrafilter.

It follows by the Ultrafilter Theorem that Δ can be extended to an ultrafilter U', and it is easy to see that U' has the desired properties. ■

Finally, closure property of the ultrafilter extensions goes in the other direction: If K is modally definable and $\mathfrak{ue}\mathfrak{F} \in \mathsf{K}$, then $\mathfrak{F} \in \mathsf{K}$. We describe this by saying that a modally definable class K of frames **reflects ultrafilter extensions**.

Consider the class of frames $\mathsf{K} = \{\mathfrak{F} \mid \mathfrak{F} \models \forall x \exists y (Rxy \wedge Ryy)\}$. That is, the class of frames which have the property that every point has a reflexive successor. This class is closed under generated subframes, disjoint unions, and bounded morphic images, but it does not reflect ultrafilter extensions. So this class of frames is modally undefinable.

To show why $\mathsf{K} = \{\mathfrak{F} \mid \mathfrak{F} \models \forall x \exists y (Rxy \wedge Ryy)\}$ fails to reflect ultrafilter extensions, let $\mathfrak{F} = (\mathbb{N}, <)$ be the frame based on the natural numbers with the usual strict ordering. The ultrafilter extension $\mathfrak{ue}\mathfrak{F}$ has a submodel that is isomorphic to $\mathfrak{F}$ (namely, the submodel consisting of the principal ultrafilters generated by the natural numbers), but there are also many ultrafilters that form clusters after the natural numbers. More precisely, we claim that for every $U \in \mathfrak{ue}\mathfrak{F}$ which is non-principal, we have RU_1U for every $U_1 \in \mathfrak{ue}\mathfrak{F}$. To see this, let $U_1 \in \mathfrak{ue}\mathfrak{F}$. We show that $\{m_R(X) \mid X \in U\} \subseteq U_1$. $X \in U$ implies that X is infinite, which implies that $m_R(X) = \mathbb{N}$. Thus, $m_R(X) \in U_1$. The above reasoning shows that $\mathfrak{ue}\mathfrak{F} \in \mathsf{K}$. However, $\mathfrak{F} \notin \mathsf{K}$, since for no $n \in \mathbb{N}$ do we have $n < n$.

Also, it is not true in general that modally definable classes of frames are closed under taking ultrafilter extensions (i.e., $\mathfrak{F} \in \mathsf{K}$ does not imply that $\mathfrak{ue}\mathfrak{F} \in \mathsf{K}$). For an example, consider the frame $\mathfrak{F} = (\mathbb{Z}^-, <)$ consisting of the negative integers strictly ordered in the usual way. This frame validates Löb's formula because $<$ is transitive and conversely well-founded. However, $\mathfrak{ue}\mathfrak{F}$ does not validate Löb's formula, as we can show that if U is a non-principal ultrafilter, then RUU, and so there is an infinite ascending chain $URURU \ldots$. That is, going from $\mathfrak{F}$ to $\mathfrak{ue}\mathfrak{F}$ does not preserve the validity of modal formulas. However, under certain conditions, modally definable classes are closed under taking ultrafilter extensions. For instance, if K is both modally definable and first order definable, then K is closed under taking ultrafilter extensions. We will prove something related to this in the proof of the Goldblatt-Thomason theorem.

9.5 Preservation Results

We have shown examples of first order formulas that do not have the above four preservation properties (preserved under taking generated subframes, disjoint unions, bounded morphic images, and ultrafilter extensions). It is natural to ask the question: Which first order formulas have the above preservation properties?

We know which first order formulas are closed under generated subframes. Feferman and Kreisel (1966) considered this question, and they show that the following syntactic shape is necessary and sufficient for preservation under generated subframes:

$$Rxy \mid \neg \mid \vee \mid \forall x \mid \exists y(Rxy \wedge \varphi).$$

In van Benthem's 1977 thesis [13], there are results for the first three preservation properties and their combination. It is still an open problem, however, to specify a syntactic criterion for preservation under ultrafilter extensions (and under all the four preservation properties combined).

In fact, this may not even be an RE class of first-order formulas. Is it the case that, given any operation on the models, we can prove a syntactic preservation theorem for it? This seems a naive expectation. If we have a syntactic preservation result, this means there is a well-defined, recursive class of syntactic shapes, and the formulas is equivalent to something in that class of shapes. The complexity of this is Σ_1^0 (RE). This is very low down in the complexity hierarchy.

Accordingly, the operation on the frames must be simple in order to have a syntactic preservation theorem, and the first three operations on frames are indeed simple. For instance, it can be shown that the first order formulas that are preserved under generated subframes must be Σ_1^0, because the claim that the truth of a first order formula φ is preserved under generated subframes is equivalent to the validity of the following first order formula:

$$\forall x(Ax \to \forall y(Rxy \to Ay) \to (\varphi \to (\varphi)^A)$$

The problem with the ultrafilter extension is that the complexity of this operation is unclear. So it is possible that there is no syntactic preservation theorem for all the four conditions.

10 Background in Universal Algebra

Another important perspective on modal logic is the algebraic perspective. The algebraic treatment of modal logic is an extension of the algebraic treatment of classical propositional logic, and it allows us to bring algebraic techniques to bear on certain model-theoretic issues. As an illustration, we will give an algebraic proof of the Goldblatt-Thomason Theorem, and the purpose of this section is to introduce some basic algebraic concepts and results that will be useful in that proof. We first extend boolean algebras to boolean algebras with operators, and then prove the Jónsson-Tarski Representation Theorem, which is an extension of the Stone Representation Theorem. Finally we briefly mention a well-known result in universal algebra, namely Birkhoff's Variety Theorem.

10.1 Boolean Algebras With Operators (BAO)

The algebraic treatment of classical propositional logic makes use of *boolean algebras*:

Definition 10.1. ***(Boolean Algebras).*** *A structure* $\mathfrak{A} = (A, 0, +, -)$ *is called a* ***boolean algebra*** *iff it satisfies the following equations:* ($x \cdot y$ *and* 1 *are shorthand for* $-(-x + -y)$ *and* -0*, respectively.*)

1. *Associativity. For both* $+$ *and* $\cdot$ *.*
2. *Commutativity. For both* $+$ *and* $\cdot$ *.*
3. *Distributivity of* $+$ *over* $\cdot$ *and vice versa.*
4. *Complementation.* $x + (-x) = 1$ *and* $x \cdot (-x) = 0$*.*
5. *Identity.* $x + 0 = x$ *and* $x \cdot 1 = x$*.*

The set A *is called the carrier set of* $\mathfrak{A}$*, and the operations* $+$ *and* $\cdot$ *are called join and meet, respectively. Moreover, we order the elements of* $\mathfrak{A}$ *by defining* $a \leq b$ *if* $a + b = b$ (*or equivalently, if* $a \cdot b = a$)*.*

The intuitive semantics of propositional logic, for instance, can be regarded as a boolean algebra: We can think of 0 as False, 1 as True, $A = \mathbf{2}$ as the set of truth values $\{0, 1\}$, and the three operations on A as operations on truth values. It is straightforward to provide translation schemes from logical formulas to algebraic equations (and vice versa), and the laws of propositional logic can then be regarded as algebraic equations that are true in the algebra of **2**. For instance, instead of saying that $p \vee \neg p$ is a law in propositional logic, we can say that the equation $x + \bar{x} = 1$ is true in the algebra of **2**.

Two kinds of Boolean algebras are particularly important for the algebraic proof of the completeness of classical propositional logic. *Set algebras* are useful for characterizing the semantics of propositional logic, whereas *Lindenbaum algebras* are useful for characterizing the syntax of propositional logic.

Definition 10.2. ***(Set Algebras)*** *Let A be a set. The* ***power set algebra*** *on A is the structure*

$$\mathfrak{A} = (\mathcal{P}(A), \emptyset, \cup, -)$$

where $\emptyset$ denotes the emptyset, $-$ is the set complement operation, and $\cup$ is the set union operation. A ***set algebra*** *is a subalgebra of a power set algebra. That is, a set algebra (on A) is a collection of subsets of A that contains $\emptyset$ and are closed under $\cup$ and $-$.*

It is straightforward to check that set algebras are boolean algebras. For an example of a set algebra that is not also a power set algebra, take all the finite and cofinite subsets of $\mathbb{N}$. It is closed under complement, union, intersection, and it includes the empty set and the whole set.

Definition 10.3. ***(Lindenbaum Algebra)*** *Let $\mathcal{L}$ be a language and $\equiv$ be the relation of provable equivalence in $\mathcal{L}$. A Lindenbaum algebra $\mathfrak{L}$ on $\mathcal{L}$ is the structure $(Form(\mathcal{L})/\equiv, 0, +, -)$, where $Form(\mathcal{L})/\equiv$ is the set of formulas in $\mathcal{L}$ modulo $\equiv$, $0 = [\bot]$, $[\varphi] + [\psi]$ is $[\varphi \vee \psi]$, and $-[\varphi]$ is $[\neg\varphi]$.*

It is straightforward to show that all these operations are well defined, and the equations that are satisfied in the Lindenbaum algebra are merely (translations of) the laws in the proof systems.

Just as boolean algebras are the key to the algebraization of classical propositional logic, in modal logic we are interested in *boolean algebra with operators* (we only need one operator here since we are focused on the basic modal language):

Definition 10.4. ***(Boolean Algebras With Operators).*** *A structure $\mathfrak{A} = (A, 0, +, -, f)$ is a* ***boolean algebra with operators*** *if $(A, 0, +, -)$ is a boolean algebra and f is a unary operation on A satisfying the following equations:*

1. *Normality.* $f(0) = 0$.

2. *Additivity.* $f(x + y) = f(x) + f(y)$.

The two equations correspond to the modal formulas $\Diamond\bot \equiv \bot$ and $\Diamond(p \vee q) \leftrightarrow (\Diamond p \vee \Diamond q)$ respectively, both of which are valid in normal modal logics.

A type of BAO is particularly important for capturing the semantics of modal logic, namely *modal algebras*. A modal algebra is basically a set algebra augmented with an operation corresponding to the accessibility relation in the frame:

Definition 10.5. ***(Modal Algebras)****. Let $\mathfrak{F} = (W, R)$ be a frame. The* ***full modal algebra*** *of $\mathfrak{F}$ (notation: $\mathfrak{F}^+$), is the structure*

$$\mathfrak{F}^+ = (\mathcal{P}(W), \cup, -, \emptyset, m_R)$$

Where m_R is the unary operation such that for any $X \subseteq W$,

$$m_R(X) = \{y \in W \mid \textit{there is an } x \in X \textit{ such that } Ryx\}$$

A ***modal algebra*** *is a subalgebra of a full modal algebra.*

As an example of modal algebra, consider again the collection of all the finite and cofinite subsets of $(\mathbb{N}, <)$ (think of $(\mathbb{N}, <)$ as a frame). This structure is closed under the operation $m_<$: If we have a finite set X, $m_<(X)$ is finite. If X is a co-finite set, $m_<(X)$ have to be the entire $\mathbb{N}$, since there is no $<$-upper bound on the co-finite set.

10.2 Jónsson-Tarski Theorem

The Stone Representation Theorem says that any boolean algebra is isomorphic to a set algebra. As a generalization of the Stone Representation Theorem, the Jónsson-Tarski Theorem says that any boolean algebra with operators is isomorphic to a modal algebra. To prove this theorem, we first construct a frame from a BAO by forming the *ultrafilter frame*. By turning this frame back into a full modal algebra, we can then define a canonical isomorphic embedding of the original BAO into this full modal algebra.

Definition 10.6. ***(Ultrafilters in Boolean Algebras).*** *A **filter** of a boolean algebra $\mathfrak{A} = (A, 0, +, -)$ is a subset $F \subseteq A$ satisfying:*

1. $1 \in F$
2. *If $a, b \in F$. then $a \cdot b \in F$.*
3. *If $a \in F$ and $a \leq b$, then $b \in F$.*

*A filter of $\mathfrak{A}$ F is **proper** iff $0 \notin F$. An **ultrafilter** of $\mathfrak{A}$ is a proper filter F such that for every $a \in A$, either a or $-a$ belongs to F. The collection of ultrafilters of $\mathfrak{A}$ is written as $Uf\mathfrak{A}$.*

Theorem 10.7. ***(Ultrafilter Theorem).*** *Let $\mathfrak{A} = (A, 0, +, -)$ be a boolean algebra.*

1. *Any proper filter of $\mathfrak{A}$ can be extended into an ultrafilter of $\mathfrak{A}$.*
2. *Suppose D is a subset of A that has the **finite meet property**; that is, no finite subset $\{a_0, \ldots, a_n\}$ of D has $a_0 \cdot \cdots \cdot a_n = 0$. Then D can be extended to an ultrafilter of $\mathfrak{A}$.*

Definition 10.8. ***(Ultrafilter Frames).*** *Let $\mathfrak{A}$ be a boolean algebra with operator f. The **Ultrafilter frame** of $\mathfrak{A}$ (notation: $\mathfrak{A}_+$) is the structure*

$$\mathfrak{A}_+ = (Uf\mathfrak{A}, R_f)$$

where $Uf\mathfrak{A}$ is the set of ultrafilters of $\mathfrak{A}$, and for any ultrafilters $u, v \in Uf\mathfrak{A}$, $R_f uv$ iff $fa \in u$ for all $a \in v$. Alternatively, if we use $f^*(a)$ as a shorthand for $-f(-a)$, then we have that $R_f uv$ iff $a \in v$ for all $f^*(a) \in u$.

Theorem 10.9. ***(Stone Representation Theorem).*** *Let $\mathfrak{A} = (A, 0, +, -)$ be a boolean algebra. Then the representation function $\rho : A \to \mathcal{P}(Uf\mathfrak{A})$ given by*

$$\rho(a) = \{u \in Uf\mathfrak{A} \mid a \in u\}$$

is an embedding of $\mathfrak{A}$ into the powerset of $Uf\mathfrak{A}$.

Proof. See Bell and Slomson's well-known textbook "Models and Ultraproducts", p.24. ■

Theorem 10.10. ***(Jónsson-Tarski Theorem).*** *Let $\mathfrak{A} = (A, 0, +, -, f)$ be a boolean algebra with operator. Then the representation function $\rho : A \to \mathcal{P}(Uf\mathfrak{A})$ given by*

$$\rho(a) = \{u \in Uf\mathfrak{A} \mid a \in u\}$$

is an embedding of $\mathfrak{A}$ into $(\mathfrak{A}_+)^+$.

Proof. By the Stone Representation Theorem, ρ is a well-defined, injective mapping, and it preserves the boolean operations $0, +$, and $-$. The only additional thing to check is that ρ preserves the modal operation f, that is, for every $a \in A$

$$\rho(f(a)) = m_{R_f}(\rho(a))$$

For the inclusion from right to left: Suppose $u \in m_{R_f}(\rho(a))$. Then by the definition of m_{R_f}, there exists an ultrafilter u_1 with $u_1 \in \rho(a)$ (that is, $a \in u_1$) and $R_f uu_1$. By the definition of R_f, $f(a) \in u$, that is $u \in \rho(f(a))$.

For the inclusion from left to right: let $u \in \rho(f(a))$, then $f(a) \in u$. To show that $u \in m_{R_f}(\rho(a))$, we need to find an ultrafilter u_1 with $a \in u_1$ and $R_f uu_1$, and for this, it suffices to show that there is some ultrafilter u_1 such that $\{a\} \cup \{x \mid f^*(x) \in u\} \subseteq u_1$, with f^* being a shorthand for $-f(-x)$. To find such a u_1 we use the ultrafilter theorem: and to set this up, we show that $\{a\} \cup \{x \mid f^*(x) \in u\}$ has the finite meet property.

First, we show that $F = \{x \mid f^*(x) \in u\}$ is closed under taking meet. Suppose $x, y \in F$, that is, $f^*(x)$, $f^*(y) \in u$. Since u is an ultrafilter, $f^*(x) \cdot f^*(y) \in u$. But the additivity axiom $f(x + y) = f(x) + f(y)$ yields $f^*(x \cdot y) = f^*(x) \cdot f^*(y)$, which in turn implies that $f^*(x \cdot y) \in u$. Hence $x \cdot y \in F$.

Next, as F is closed under taking meets, to show that $\{a\} \cup F$ has finite meet property, it suffices to show that $a \cdot x \neq 0$ whenever $x \in F$. To arrive at a contradiction, suppose that $a \cdot x = 0$ for $x \in F$. Then $a = a \cdot 1 = a \cdot (x + (-x)) = (a \cdot x) + (a \cdot (-x)) = 0 + (a \cdot (-x)) = a \cdot (-x)$, which implies that $a \leq -x$. It is easy to show that f is monotonic, hence $f(a) \leq f(-x)$. Since $f(a) \in u$, we have $f(-x) \in u$ and so $f^*(x) \notin u$, contradicting $x \in F$. ■

The Jónsson-Tarski Theorem tells us what happens if we start with a BAO, apply the ultrafilter frame construction and then the modal algebra construction. What will happen if we start with a frame, construct the modal algebra of this frame, and then construct the ultrafilter frame of this modal algebra? The following result gives the answer:

Theorem 10.11. *Let $\mathfrak{F}$ be a frame. Then $(\mathfrak{F}^+)_+$, the ultrafilter frame of the modal algebra of $\mathfrak{F}$, is the ultrafilter extension $\mathfrak{ue}\mathfrak{F}$ of $\mathfrak{F}$.*

10.3 Birkhoff's Theorem

A fundamental result in universal algebra is Birkhoff's Variety Theorem, which says that a collection of algebras is equationally definable if and only if it is closed under three algebraic operations. In this section, we briefly review the terminologies needed for formulating Birkhoff's Theorem:

Recall that a *language* (or *signature*) of algebras is a set $\mathcal{F}$ of functional symbols such that a nonnegative integer n is assigned to each member $f \in \mathcal{F}$. This integer is called the *arity* of f, and f is said to be an n-ary function symbol. The subset of n-ary function symbols in $\mathcal{F}$ is denoted by $\mathcal{F}_n$. Given a language $\mathcal{F}$, an *algebra* $\mathfrak{A} = (A, f^{\mathfrak{A}})_{f \in \mathcal{F}}$ of signature $\mathcal{F}$ consists of a set A and an operation $f^{\mathfrak{A}}$ on A corresponding to each function symbol f in $\mathcal{F}$.

Let X be a set of variables and $\mathcal{F}$ be a signature of algebras. The set T of *terms* of signature $\mathcal{F}$ is the smallest set such that

1. $X \cup \mathcal{F}_0 \subseteq T$
2. If $t_1, \ldots, t_n \in T$ and $f \in \mathcal{F}_n$, then $f(t_1, \ldots t_n) \in T$.

An *equation* of signature $\mathcal{F}$ over a set of variables X is an expression of the form $t_1 \approx t_2$, where $t_1, t_2 \in T$. An algebra $\mathfrak{A}$ of type $\mathcal{F}$ *satisfies* an equation $t_1(x_1, \ldots x_n) \approx t_2(x_1 \ldots x_n)$, abbreviated by $\mathfrak{A} \models t_1(x_1, \ldots x_n) \approx t_2(x_1 \ldots x_n)$, if for every choice of $a_1, \ldots a_n \in A$ we have $t_1^{\mathfrak{A}}(a_1, \ldots, a_n) = t_2^{\mathfrak{A}}(a_1, \ldots, a_n)$. A class K of algebras satisfies $t_1 \approx t_2$ if every each member of K satisfies $t_1 \approx t_2$.

Definition 10.12. ***(Equational Class)***. *Let Σ be a set of equations of signature $\mathcal{F}$, and define* $\mathrm{Mod}(\Sigma)$ *to be the class of algebras satisfying Σ. A class K of algebras is an equational class (also:*

'equational variety') if there is a set of equations Σ such that $\mathsf{K} = \mathrm{Mod}(\Sigma)$. In this case we also say that K is equationally defined by Σ.

Definition 10.13. ***(Homomorphism)****. Let $\mathcal{F}$ be an algebraic signature, and let $\mathfrak{A} = (A, f^{\mathfrak{A}})_{f\in\mathcal{F}}$ and $\mathfrak{B} = (B, f^{\mathfrak{B}})_{f\in\mathcal{F}}$ be two algebras of the same signature. a map $\eta : A \to B$ is a* ***homomorphism*** *iff for all $f \in F$ and all $a_1, \ldots, a_n \in A$:*

$$\eta(f^{\mathfrak{A}}(a_1, \ldots, a_n) = f^{\mathfrak{B}}(\eta(a_1), \ldots, \eta(a_n))$$

Definition 10.14. ***(Subalgebra)****. Let $\mathfrak{A}$ be an algebra with signature $\mathcal{F}$ and B be a subset of the carrier A. If B is closed under every operation $f_{\mathfrak{A}}$, then we call $\mathfrak{B} = (B, f_{\mathfrak{A}} \restriction_B)_{f\in F}$ a* ***subalgebra*** *of $\mathfrak{B}$.*

Definition 10.15. ***(Product)****. Let $(\mathfrak{A}_i)_{i\in I}$ be a family of algebras. The* ***product*** *$\prod_{i\in I}\mathfrak{A}_i$ of the family is the algebra $\mathfrak{A} = (A, f^{\mathfrak{A}})_{f\in F}$ where A is the cartesian product $\prod_{i\in I} A_i$ for each carrier A_i of $\mathfrak{A}_i$, and the operations are defined coordinatewise; that is, for functions $a_1, \ldots, a_n \in \prod_{i\in I} A_i$, $f^{\mathfrak{A}}(a_1, \ldots, a_n)$ is the function in $\prod_{i\in I} A_i$ given by:*

$$f^{\mathfrak{A}}(a_1, \ldots, a_n)(i) = f_i^{\mathfrak{A}}(a_1(i), \ldots, a_n(i))$$

Theorem 10.16. ***(Birkhoff's Theorem)*** *A class of algebras K (in any signature) is an equational class if and only if K is closed under taking homomorphic image, subalgebras, and products.*

In fact, Birkhoff's theorem shows how to generate an equational class from a given class of algebras: Given a class C of algebras, let $\mathbb{V}\mathsf{C}$ denote the equational class generated by C. That is, let $\Sigma(\mathsf{C})$ be the set of all the equations that are true in C, and $\mathbb{V}\mathsf{C}$ is the class of all the algebras satisfying $\Sigma(\mathsf{C})$. Birkhoff's theorem (together with a theorem by Tarski) shows that $\mathbb{V}\mathsf{C} = \mathbf{HSP}\mathsf{C}$. That is, in order to obtain the equational class generated by C, we can start by taking products of algebras in C, then go on to take subalgebras, and finish off by forming homomorphic images.

11 Modal Duality Theory

So far we have seen how to construct frames from algebras and algebras from frames. But frames can be constructed from other frames by generated subframes, disjoint unions and bounded morphisms, and algebras can be constructed from other algebras by homomorphisms, subalgebras, and products. Modal duality theory is the study of the systematic connection between these frame-building operations and algebra-building operations. The duality results we prove in this section focus only on the direction from algebras to frames, because this is the direction needed in the algebraic proof of the Goldblatt-Thomason Theorem.

Theorem 11.1. *Let $\mathfrak{F}_i = (W_i, R_i)$, $i \in I$ be a family of frames. Then the modal algebra of the disjoint union of $\mathfrak{F}_i$ $(i \in I)$ is isomorphic to the the product of the modal algebras of $\mathfrak{F}_i$ $(i \in I)$.*

$$\left(\biguplus_{i\in I} \mathfrak{F}_i\right)^+ \cong \prod_{i\in I} \mathfrak{F}_i^+$$

Proof. (Sketch only.) We define a map $\nu : \mathcal{P}(\biguplus_{i\in I} \mathfrak{F}_i) \longrightarrow \prod_{i\in I} \mathcal{P}(W_i)$ such that, for any $X \subseteq \biguplus_{i\in I} \mathfrak{F}_i$ and any $i \in I$,

$$\nu(X)(i) = X \cap W_i$$

It is easy to see that ν is an isomorphism. ■

Definition 11.2. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be two boolean algebras with operators, and $\eta : A \to B$ be a function from the carrier of $\mathfrak{A}$ to the carrier of $\mathfrak{B}$. Then its* ***dual*** *η_+ is the following function that maps each ultrafilter u of $\mathfrak{B}$ to a subset of A:*

$$\eta_+(u) = \{a \in A \mid \eta(a) \in u\}$$

Definition 11.3. *Let $\mathfrak{A} = (A, 0^{\mathfrak{A}}, +^{\mathfrak{A}}, -^{\mathfrak{A}}, f^{\mathfrak{A}})$ and $\mathfrak{B} = (B, 0^{\mathfrak{B}}, +^{\mathfrak{B}}, -^{\mathfrak{B}}, f^{\mathfrak{B}})$ be two booleans algebras with operator. and let $\eta : A \to B$ be a function. We say that η is a* ***boolean homomorphism*** *if η is a homomorphism if η is a homomorphism from $(A, 0^{\mathfrak{A}}, +^{\mathfrak{A}}, -^{\mathfrak{A}})$ to $(B, 0^{\mathfrak{B}}, +^{\mathfrak{B}}, -^{\mathfrak{B}})$. We say that η is a* ***modal homomorphism*** *if for all $a \in A$, $\eta(f^A(a)) = f^{\mathfrak{B}}(\eta(a))$. Finally, η is a* ***homomorphism*** *if it is both a boolean and a modal homomorphism.*

Lemma 11.4. *Let $\mathfrak{A} = (A, 0^{\mathfrak{A}}, +^{\mathfrak{A}}, -^{\mathfrak{A}}, f^{\mathfrak{A}})$ and $\mathfrak{B} = (B, 0^{\mathfrak{B}}, +^{\mathfrak{B}}, -^{\mathfrak{B}}, f^{\mathfrak{B}})$ be two booleans algebras with operator, and η a map from A to B.*

1. *If η is a boolean homomorphism, then η_+ maps ultrafilters to ultrafilters.*
2. *If $f^{\mathfrak{B}}(\eta(a)) \leq^{\mathfrak{B}} \eta(f^{\mathfrak{A}}(a))$, then η_+ has the zig property of bounded morphism.*
3. *If $f^{\mathfrak{B}}(\eta(a)) \geq^{\mathfrak{B}} \eta(f^{\mathfrak{A}}(a))$ and η is a boolean homomorphism, then η_+ has the zag property of bounded morphism.*
4. *If η is a homomorphism, then η_+ is a bounded morphism from $\mathfrak{B}_+$ to $\mathfrak{A}_+$.*
5. *If η is an injective boolean homomorphism, then $\eta_+ : Uf\mathfrak{B} \to Uf\mathfrak{A}$ is surjective.*
6. *If η is a surjective boolean homomorphism, then $\eta_+ : Uf\mathfrak{B} \to Uf\mathfrak{A}$ is injective.*

Proof. 1. Suppose η is a boolean homomorphism. We show that $\eta_+(u) = \{a \in A \mid \eta(a) \in u\}$ is an ultrafilter of $\mathfrak{A}$. First, we show that $0^{\mathfrak{A}} \notin \eta_+(u)$. Since η is a boolean homomorphism and u is an ultrafilter of $\mathfrak{B}$, $\eta(0^{\mathfrak{A}}) = 0^{\mathfrak{B}} \notin u$. Hence $0^{\mathfrak{A}} \notin \eta_+(u)$. Second, suppose $a, b \in \eta_+(u)$, that is, $\eta(a), \eta(b) \in u$. We show that $a \cdot^{\mathfrak{A}} b \in \eta_+(u)$. Again since η is a boolean homomorphism and u is an ultrafilter of $\mathfrak{B}$, $\eta(a \cdot^{\mathfrak{A}} b) = \eta(a) \cdot^{\mathfrak{B}} \eta(b) \in u$, and so $a \cdot^{\mathfrak{A}} b \in \eta_+(u)$. Similarly, it is straightforward to show that $\eta_+(u)$ is upward closed and that for every $a \in A$, either a or $-^{\mathfrak{A}}a$ belongs to $\eta_+(u)$.

2. We need to show that for any $u, u_1 \in Uf\mathfrak{B}$, if $R_{f^{\mathfrak{B}}}uu_1$, then $R_{f^{\mathfrak{A}}}\eta_+(u)\eta_+(u_1)$. Let $a \in \eta_+(u_1)$, that is, $\eta(a) \in u_1$. Then $f^{\mathfrak{B}}(\eta(a)) \in u$ since $R_{f^{\mathfrak{B}}}uu_1$, and so $\eta(f^{\mathfrak{A}}(a)) \in u$ by the assumption and by the fact that u is upward closed. Hence $f^{\mathfrak{A}}(a) \in \eta_+(u)$ as desired.

3. Suppose $R_{f^{\mathfrak{A}}}\eta_+(u)v_1$ holds for $u \in Uf\mathfrak{B}$ and $v_1 \in Uf\mathfrak{A}$. We want to find $u_1 \in Uf\mathfrak{B}$ such that $\eta_+(u_1) = v_1$ and $R_{f^{\mathfrak{B}}}uu_1$. Define

$$F = \{\eta(a) \mid a \in v_1\} \cup \{b \in B \mid -^{\mathfrak{B}} f^{\mathfrak{B}}(-^{\mathfrak{B}}b) \in u\}$$

We show that F has the finite meet property. It is easy to see that both $\{\eta(a) \mid a \in v_1\}$ and $\{b \in B \mid -^{\mathfrak{B}} f^{\mathfrak{B}}(-^{\mathfrak{B}}b) \in u\}$ are closed under meet, hence it suffices to show that $\eta(a) \cdot^{\mathfrak{B}} b \neq 0^{\mathfrak{B}}$ for $a \in v_1$ and $b \in B$. Suppose for contradiction that $\eta(a) \cdot^{\mathfrak{B}} b = 0^{\mathfrak{B}}$. Then $\eta(a) \leq^{\mathfrak{B}} -^{\mathfrak{B}}b$, and by the monotonicity of $f^{\mathfrak{B}}$, $f^{\mathfrak{B}}(\eta(a)) \leq^{\mathfrak{B}} f^{\mathfrak{B}}(-^{\mathfrak{B}}b)$. By assumption, $\eta(f^{\mathfrak{A}}(a)) \leq^{\mathfrak{B}} f^{\mathfrak{B}}(\eta(a))$. Moreover,

$\eta(f^{\mathfrak{A}}(a)) \in u$, since $a \in v_1$ and $R_{f^{\mathfrak{A}}}\eta_+(u)v_1$. Hence by the upward closure of u, $f^{\mathfrak{B}}(-^{\mathfrak{B}}b) \in u$, contradicting the assumption that $-^{\mathfrak{B}}f^{\mathfrak{B}}(-^{\mathfrak{B}}b) \in u$.

By the ultrafilter theorem, the filter F can be extended to an ultrafilter u_1, and we will now show that u_1 in fact has the desired properties. First, $R_{f^{\mathfrak{B}}}uu_1$ because $\{b \in B \mid -^{\mathfrak{B}} f^{\mathfrak{B}}(-^{\mathfrak{B}}b) \in u\} \subseteq F \subseteq u_1$. Second, to show that $\eta_+(u_1) = v_1$, first let $a \in v_1$, then $\eta(a) \in \{\eta(a) \mid a \in v_1\} \subseteq u_1$. This shows that $v_1 \subseteq \eta_+(u_1)$. For the other inclusion, it suffices to show that $a \notin \eta_+(u_1)$ if $a \notin v_1$; we reason as follows:

$$
\begin{aligned}
a \notin v_1 \quad &\Longrightarrow \quad -^{\mathfrak{A}}a \in v_1 \\
&\Longrightarrow \quad -^{\mathfrak{B}}(\eta(a)) = \eta(-^{\mathfrak{A}}(a)) \in \{\eta(a) \mid a \in v_1\} \\
&\Longrightarrow \quad -^{\mathfrak{B}}(\eta(a)) \in u_1 \\
&\Longrightarrow \quad \eta(a) \notin u_1 \\
&\Longrightarrow \quad a \notin \eta_+(u_1)
\end{aligned}
$$

4. This follows immediately from item 1, 2 and 3.

5. This proof is similar to the proof of 3. Assume that η is injective, and let v be an ultrafilter of $\mathfrak{A}$. we want to find an ultrafilter u of $\mathfrak{B}$ such that $\eta_+(u) = v$. Define

$$\eta[v] = \{\eta(a) \mid a \in u\}$$

It is straightforward to show that $\eta[v]$ has the finite meet property and therefore can be extended to an ultrafilter u. We can then show that $\eta_+(u) = v$ by following the same reasoning as in part 3.

6. Assume that η is surjective, and let u_1, u_2 be distinct ultrafilters of $\mathfrak{A}$. Without loss of generality we may assume that there exists an $b \in B$ such that $a \in u_1$ and $a \notin u_2$. As η is surjective, there is an $a \in A$ such that $\eta(a) = b$. Then $a \in \eta_+(u_1)$ but $a \notin \eta_+(u_2)$, and so $\eta_+(u_1) \neq \eta_+(u_2)$. Hence η_+ is injective. ∎

Theorem 11.5. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be two boolean algebras with operators.*

1. *If $\mathfrak{A}$ is isomorphic to a subalgebra of $\mathfrak{B}$, then $\mathfrak{A}_+$ is a bounded morphic image of $\mathfrak{B}_+$.*

2. *If $\mathfrak{A}$ is a homomorphic image of $\mathfrak{B}$, then $\mathfrak{A}_+$ is isomorphic to a generated subframe of $\mathfrak{B}_+$.*

Proof. This is an immediate consequence of Lemma 11.4. ■

12 Goldblatt-Thomason Theorem

In this section we provide two proofs of the Goldblatt-Thomason Theorem: The first proof is algebraic whereas the second is model-theoretic.

12.1 The First Proof

Lemma 12.1. *Let $\mathfrak{F} = (W, R)$ be an arbitrary frame. Then $\mathfrak{F}$ has an ultrapower $\prod_U \mathfrak{F}$ such that $\prod_U \mathfrak{F}$ is an elementary extension of $\mathfrak{F}$, and the ultrafilter extension $\mathfrak{ue}(\mathfrak{F})$ is the bounded morphic image of $\prod_U \mathfrak{F}$.*

Proof. Let $\Sigma = \{p_A \mid A \subseteq W\}$. That is, Σ is a set of new proposition letters, one for each subset of W. Define a valuation V such that $V(p_A) = A$ for each $A \subseteq W$, and let $\mathfrak{M} = (\mathfrak{F}, V)$. By Theorem 2.16, there is a $\omega-$saturated ultrapower $\prod_U \mathfrak{M}$ which is an elementary extension of $\mathfrak{M}$. (We cannot apply corollary 2.13 here since Σ may not be countable.) Now we define a mapping $f : \prod_U \mathfrak{M} \to \mathfrak{ue}\mathfrak{M}$:

$$\text{For every } s \in \prod_U \mathfrak{M},\ f(s) = \{A \subseteq W \mid \prod_U \mathfrak{M}, s \vDash p_A\}$$

and we show that f is a surjective bounded morphism.

First, we show that f is well defined in the sense that $f(s)$ is indeed an ultrafilter. We will only check that $f(s)$ does not contain the empty set and is closed under intersection, since the other conditions are similar. By the definition of f, we have $\emptyset \in f(s)$ iff $\prod_U \mathfrak{M}, s \vDash p_\emptyset$. By the definition of V, we have $\mathfrak{M} \vDash \neg p_\emptyset$. Note that $\mathfrak{M}$ and $\prod_U \mathfrak{M}$ are elementarily equivalent, which implies that for all modal formulas φ, $\mathfrak{M} \vDash \varphi$ iff $\prod_U \mathfrak{M} \vDash \varphi$. Hence we have $\prod_U \mathfrak{M} \vDash \neg p_\emptyset$, and in particular $\prod_U \mathfrak{M}, s \vDash \neg p_\emptyset$. Hence $\emptyset \notin f(s)$. Similarly, let $X, Y \in f(s)$, that is, $\prod_U \mathfrak{M}, s \vDash p_X \wedge p_Y$. By the definition of V, we have $\mathfrak{M} \vDash (p_X \wedge p_Y) \leftrightarrow p_{X\cap Y}$. Hence $\prod_U \mathfrak{M} \vDash (p_X \wedge p_Y) \leftrightarrow p_{X\cap Y}$, and so $\prod_U \mathfrak{M}, s \vDash p_{X\cap Y}$. Then $X \cap Y \in f(s)$ as desired.

Second, we show that f is homomorphic (the 'zig' clause of bounded morphism). Suppose $s_1, s_2 \in \prod_U \mathfrak{M}$ and $R^U s_1 s_2$. We want to show that $R^{\mathfrak{ue}} f(s_1) f(s_2)$; that is, if $X \in f(s_2)$, then

$m(X) \in f(s_1)$. Suppose $X \in f(s_2)$, hence $\prod_U \mathfrak{M}, s_2 \models p_X$, and so $\prod_U \mathfrak{M}, s_1 \models \Diamond p_X$. Since $\mathfrak{M} \models \Diamond p_X \leftrightarrow p_{m(X)}$, we have $\prod_U \mathfrak{M} \models \Diamond p_X \leftrightarrow p_{m(X)}$. Therefore $\prod_U \mathfrak{M}, s_1 \models p_{m(X)}$, and so $m(X) \in f(s_1)$ as desired.

Third, to show the 'zag' clause of bounded morphism, it suffices to show that $\{(s, f(s)) \mid s \in \prod_U \mathfrak{M}\}$ is a bisimulation. Now, both $\prod_U \mathfrak{M}$ and $\mathfrak{ue}\mathfrak{M}$ are modally saturated. By Lemma 3.5, it suffices to show that for any $s \in \prod_U \mathfrak{M}$ and $u \in \mathfrak{ue}\mathfrak{M}$, $u = f(s)$ if and only if s and u are modally equivalent. The left-to-right direction can be proved by noting that $\mathfrak{M} \models p_{V(\varphi)} \leftrightarrow \varphi$ by the definition of V, hence $\prod_U \mathfrak{M} \models p_{V(\varphi)} \leftrightarrow \varphi$. It follows that

$$\begin{aligned} \prod_U \mathfrak{M}, s \models \varphi \ \text{ iff } & \prod_U \mathfrak{M}, s \models p_{V(\varphi)} \\ \text{iff } & V(\varphi) \in f(s) \\ \text{iff } & \mathfrak{ue}\mathfrak{M}, u \models \varphi \end{aligned}$$

To prove the right-to-left direction, Suppose that s and u are modally equivalent points. Then we have, for any set $X \subseteq W$,

$$\begin{aligned} X \in f(s) \ \text{ iff } & \prod_U \mathfrak{M}, s \models p_X \\ \text{iff } & \prod_U \mathfrak{M}, s \models p_{V(p_X)} \\ \text{iff } & \mathfrak{ue}\mathfrak{M}, u \models p_{V(p_X)} \\ \text{iff } & V(p_X) = X \in u \end{aligned}$$

Hence $f(s) = u$.

Finally, we show that f is surjective. Let $u \subseteq \mathcal{P}(W)$ be an ultrafilter over W. We want to find an object $s \in \prod_U \mathfrak{M}$ such that $f(s) = u$.

Let $\Theta = \{p_A \mid A \in u\}$. We show that Θ is finitely satisfiable in $\prod_U \mathfrak{M}$. Since u is an ultrafilter, $A_1 \cap \cdots \cap A_n \neq \emptyset$, and therefore, $p_{A_1} \wedge \cdots \wedge p_{A_n}$ is satisfiable in $\mathfrak{M}$ as well as in $\prod_U \mathfrak{M}$. By $\omega-$saturation, Θ is satisfiable in $\prod_U \mathfrak{M}$, and hence there exists $s \in \prod_U \mathfrak{M}$ such that s satisfies Θ. We claim that $f(s) = u$. Take any $X \subseteq W$. If $X \in u$, then $p_X \in \Theta$ and $\prod_U \mathfrak{M}, s \models p_X$. That is,

$X \in f(s)$. If $X \notin u$, then $\bar{X} \in u$, hence $\prod_U \mathfrak{M}, s \vDash p_{\bar{X}}$. It follows that $\bar{X} \in f(s)$, that is, $X \notin f(s)$. Hence we have $u = f(s)$, as desired. ■

Theorem 12.2. ***(Goldblatt-Thomason Theorem)*** *Let* K *be an elementary class of frames (i.e., a class of frames definable by first order formulas). Then* K *is modally definable iff* K *is closed under taking generated subframes, disjoint unions, bounded morphic images, and reflects ultrafilter extensions.*

$$
\begin{tikzcd}
\mathfrak{F} & (\mathfrak{F}^+)_+ = \mathfrak{ue}(\mathfrak{F}) \arrow[l] & \mathfrak{A}_+ \arrow[l] & & \mathfrak{B}_+ = \mathfrak{ue}(\biguplus_{i\in I}\mathfrak{F}_i) \arrow[ll] \\
& \mathfrak{F}^+ \arrow[u] & \mathfrak{A} \arrow[u] \arrow[l, "H"'] & \prod_{i\in I}\mathfrak{F}_i^+ \arrow[l, "S"'] \arrow[r, "\cong"'] & \mathfrak{B} = \left(\biguplus_{i\in I}\mathfrak{F}_i\right)^+ \arrow[u]
\end{tikzcd}
$$

Proof. ($\Longleftarrow$) Let K be any class of frames satisfying the closure conditions given in the theorem, and let Σ be the modal theory of K (that is, the set of all the modal formulas that are valid in all the frames in K.) We show that Σ defines K. By definition, we have if $\mathfrak{F} \in \mathsf{K}$ then $\mathfrak{F} \vDash \Sigma$. Thus, we just need to show that $\mathfrak{F} \vDash \Sigma$ implies $\mathfrak{F} \in \mathsf{K}$.

Suppose $\mathfrak{F} \vDash \Sigma$. We switch from $\mathfrak{F}$ to the modal algebra $\mathfrak{F}^+$ associated with $\mathfrak{F}$, and we use a translation scheme to turn the modal theory Σ of K into the equational theory Σ^+ of the class of modal algebras K^+. It is not difficult to show that the modal algebra $\mathfrak{F}^+$ satisfies Σ^+; that is, $\mathfrak{F}^+$ is in the variety generated by K^+. By Birkhoff's Theorem, $\mathfrak{F}^+ \in \mathbf{HSP}\mathsf{K}^+$. In other words, there is a family of frames $\mathfrak{F}_i \in \mathsf{K}$ ($i \in I$), and there exist a boolean algebra with operator $\mathfrak{A}$, such that $\mathfrak{A}$ is a subalgebra of the product $\prod_{i\in I}\mathfrak{F}_i^+$, and that $\mathfrak{F}^+$ is a homomorphic image of $\mathfrak{A}$.

By Theorem 11.1, $\prod_{i\in I}\mathfrak{F}_i^+$ is isomorphic to the modal algebra (call it $\mathfrak{B}$) of the disjoint union of the family of frames $(\mathfrak{F}_i)_{i\in I}$:

$$\prod_{i\in I}\mathfrak{F}_i^+ \cong \left(\biguplus_{i\in I}\mathfrak{F}_i\right)^+ = \mathfrak{B}$$

As K is closed under taking disjoint unions, $\biguplus_{i\in I}\mathfrak{F}_i$ is in K. Moreover, Theorem 10.11 implies that the modal algebra $\mathfrak{B}_+ = \left(\left(\biguplus_{i\in I}\mathfrak{F}_i\right)^+\right)_+$ is precisely the ultrafilter extension $\mathfrak{ue}(\biguplus_{i\in I}\mathfrak{F}_i)$. In order to show that $\mathfrak{B}_+ = \mathfrak{ue}(\biguplus_{i\in I}\mathfrak{F}_i)$ is in K, we need an extra closure condition not in the list of conditions of the Goldblatt-Thompson Theorem:

For any frame $\mathfrak{F}'$, if $\mathfrak{F}' \in \mathsf{K}$, then $\mathfrak{ue}(\mathfrak{F}') \in \mathsf{K}$.

This extra condition follows from Lemma 12.1: $\mathfrak{F}'$ has an ultrapower $\prod_U \mathfrak{F}'$ such that $\prod_U \mathfrak{F}'$ is an elementary extension of $\mathfrak{F}'$, and the ultrafilter extension $\mathfrak{ue}(\mathfrak{F}')$ is the bounded morphic image of $\prod_U \mathfrak{F}'$. Given that $\mathfrak{F}' \in \mathsf{K}$ and that K is an elementary class, $\prod_U \mathfrak{F}' \in \mathsf{K}$. It then follows that $\mathfrak{ue}(\mathfrak{F}') \in \mathsf{K}$, since K is closed under taking bounded morphic images.

So now we have $\mathfrak{B}_+ \in \mathsf{K}$. Recall that $\mathfrak{A}$ is isomorphic to a subalgebra of $\mathfrak{B}$, and that $\mathfrak{F}^+$ is a homomorphic image of $\mathfrak{A}$. By Theorem 11.5, $\mathfrak{A}_+$ is a bounded morphic image of $\mathfrak{B}_+$, and $(\mathfrak{F}^+)_+ = \mathfrak{ue}\mathfrak{F}$ is isomorphic to a generated subframe of $\mathfrak{A}_+$. As K is closed under taking bounded morphic images and generated subframes, it follows that $\mathfrak{ue}\mathfrak{F} \in \mathsf{K}$. But then $\mathfrak{F} \in \mathsf{K}$, since the class K reflects ultrafilter extensions. ■

A comment on the proof. The assumption that the class K is an elementary class is *only* used in proving the extra closure condition; more specifically, it is only used to justify the inference from $\mathfrak{F}' \in \mathsf{K}$ to $\prod_U \mathfrak{F}' \in \mathsf{K}$. But clearly the inference will still go through if we replace the original assumption with a purely structural assumption, namely that K is closed under ultrapowers. Since first order definability seems to be a syntactic condition, it may be preferable to replace the original syntactic assumption with a purely structural assumption.

12.2 The Second Proof

The first proof of the Goldblatt-Thomason Theorem raises a question about 'purity of methods': The theorem itself appears model-theoretic; there is nothing algebraic about it. So why wouldn't there be a model-theoretic proof for the result? Is the excursion through universal algebra really necessary? In a 1993 paper 'Modal Frame Classes Revisited', van Benthem gives a purely model-theoretic proof for the Goldblatt-Thomason Theorem. The basic strategy is to locally replace algebraic arguments in the proof by model theoretic arguments.

The proof starts with a very simple Birkhoff-like observation:

Lemma 12.3. *Any modal frame is a bounded morphic image of the disjoint union of its point-generated subframes.*

The rest of the proof is inspired by (and similar to) the model-theoretic proof of Lemma 67:

Proof. (Second Proof of the GT Theorem.)

($\Longleftarrow$) Again, we show that the modal theory Σ of K defines K, and it suffices to show that $\mathfrak{F} \vDash \Sigma$ implies $\mathfrak{F} \in \mathsf{K}$. Let $\mathfrak{F}$ be a frame such that $\mathfrak{F} \vDash \Sigma$. To show that $\mathfrak{F} \in \mathsf{K}$, it suffices to show that for every w in $\mathfrak{F}$, the subframe $\mathfrak{F}'$ of $\mathfrak{F}$ generated by the point w is in K. This is because $\mathfrak{F}$ is a bounded morphic image of the disjoint union of all its point-generated subframes, and K is closed under taking bounded morphic images and disjoint unions. So let $\mathfrak{F}' = (W, R)$ be a subframe of $\mathfrak{F}$ generated by w. Since validity is preserved under generated subframes, $\mathfrak{F}' \vDash \Sigma$.

Let $\Phi = \{p_A \mid A \subseteq W\}$. That is, Φ is a set of new proposition letters, one for each subset of W. Define a valuation V such that $V(p_A) = A$ for each $A \subseteq W$, let $\mathfrak{M} = (\mathfrak{F}', V)$ be a model for the modal language $\mathcal{L}_\Phi$, and let Δ be the collection of $\mathcal{L}_\Phi$-formulas δ such that $\mathfrak{M}, w \vDash \delta$. We claim that Δ is satisfiable in K; that is, there is a model $\mathfrak{N}'$ for $\mathcal{L}_\Phi$ and a point $v \in \mathfrak{N}'$, such that $\mathfrak{N}', v \vDash \Delta$ and the underlying frame of $\mathfrak{N}'$ is in K.

In order to show this, we first show that Δ is finitely satisfiable in K. Let Δ_0 be a finite subset of Δ, and suppose for contradiction that it is not satisfiable in K. This is equivalent to saying that for every frame $\mathfrak{F} \in \mathsf{K}$, we have $\mathfrak{F} \vDash \neg \bigwedge \Delta_0$. Since Δ_0 is finite, only finitely many proposition letters occur in it, and so $\neg \bigwedge \Delta_0$ is equivalent to a modal formula in the original modal theory Σ of K. Then we have $\neg \bigwedge \Delta_0 \in \Sigma$ by the definition of Σ. However, this contradicts the fact that $\mathfrak{M} \vDash \Sigma$ and that $\mathfrak{M}, w \vDash \delta$ for all $\delta \in \Delta$. In short, Δ is finitely satisfiable in K. Next, by Lemma 8.3, Δ is satisfiable in some ultraproduct of frames in K. Since K is an elementary class, it is closed under ultraproducts, and so Δ is satisfiable in some frame in K.

So we have a model $\mathfrak{N}'$ for $\mathcal{L}_\Phi$ and a point $v \in \mathfrak{N}$, such that $\mathfrak{N}', v \vDash \Delta$ and the underlying frame of $\mathfrak{N}'$ is in K. Now, let $\mathfrak{N}$ be the submodel of $\mathfrak{N}'$ generated by v. The underlying frame of $\mathfrak{N}'$ is still in K, since K is closed by taking generated subframes. $\mathfrak{N}, v \vDash \Delta$ since generated submodels preserve modal truth. By Theorem 2.16, there is a $\omega-$saturated ultrapower $\prod_U \mathfrak{N}$ which is an elementary extension of $\mathfrak{N}$. Since K is an elementary class, $\prod_U \mathfrak{N} \in \mathsf{K}$.

Finally we define a mapping $f : \prod_U \mathfrak{N} \to \mathfrak{ue}\mathfrak{M}$:

$$\text{For every } s \in \prod_U \mathfrak{N},\ f(s) = \{A \subseteq W \mid \prod_U \mathfrak{N}, s \vDash p_A\}$$

and we show that f is a surjective bounded morphism. If we can do this, then $\mathfrak{ue}\mathfrak{F}' \in \mathsf{K}$ since K is closed by taking bounded morphic images, and so $\mathfrak{F}' \in \mathsf{K}$ since K reflects ultrafilter extensions.

To show that f is a surjective bounded morphism, it is helpful to prove the following two facts about every modal formula φ in $\mathcal{L}_\Phi$:

1. $\mathfrak{M} \vDash \varphi$ iff $\prod_U \mathfrak{N} \vDash \varphi$

2. φ is satisfiable in $\mathfrak{M}$ implies that φ is satisfiable in $\prod_U \mathfrak{N}$

To show the first of these, we reason as follows, noting that $\mathfrak{N}$ is point-generated by v and $\mathfrak{M}$ is point-generated by w:

$$\begin{aligned}
\prod_U \mathfrak{N} \vDash \varphi \quad & \text{iff} \quad \mathfrak{N} \vDash \varphi \\
& \text{iff} \quad \mathfrak{N}, v \vDash \Box^n \varphi \text{ for all } n \in \mathbb{N} \\
& \text{iff} \quad \mathfrak{M}, w \vDash \Box^n \varphi \text{ for all } n \in \mathbb{N} \\
& \text{iff} \quad \mathfrak{M} \vDash \varphi
\end{aligned}$$

The second fact can be proved as follows:

$$\begin{aligned}
\varphi \text{ is satisfiable in } \mathfrak{M} \quad & \Longleftrightarrow \quad \mathfrak{M}, w \vDash \Diamond^N \varphi \text{ for some } n \in \mathbb{N} \\
& \Longleftrightarrow \quad \mathfrak{N}, v \vDash \Diamond^N \varphi \text{ for some } n \in \mathbb{N} \\
& \Longleftrightarrow \quad \prod_U \mathfrak{N}, v^* \vDash \Diamond^N \varphi \text{ for some } n \in \mathbb{N} \\
& \Longrightarrow \quad \varphi \text{ is satisfiable in } \prod_U \mathfrak{N}
\end{aligned}$$

Now we are ready to show that f is a surjective bounded morphism. The details of this part of the proof are basically the same as those in the proof of Lemma 12.1, except that the above two facts are used. We include the details of the proof only to show where the two facts are used.

First, we show that f is well defined in the sense that $f(s)$ is indeed an ultrafilter. We will only check that $f(s)$ does not contain the empty set. By the definition of f, we have $\emptyset \in f(s)$ iff $\prod_U \mathfrak{N}, s \vDash p_\emptyset$. By the definition of V, we have $\mathfrak{M} \vDash \neg p_\emptyset$. By fact 1, we have $\prod_U \mathfrak{N} \vDash \neg p_\emptyset$, and in particular $\prod_U \mathfrak{N}, s \vDash \neg p_\emptyset$. Hence $\emptyset \notin f(s)$ as desired.

Second, we show that f is homomorphic. Suppose $s_1, s_2 \in \prod_U \mathfrak{N}$ and $R^U s_1 s_2$. We want to show that $R^{\mathfrak{ue}} f(s_1) f(s_2)$; that is, if $X \in f(s_2)$, then $m(X) \in f(s_1)$. Suppose $X \in f(s_2)$, hence $\prod_U \mathfrak{N}, s_2 \models p_X$, and so $\prod_U \mathfrak{N}, s_1 \models \Diamond p_X$. Since $\mathfrak{M} \models \Diamond p_X \leftrightarrow p_{m(X)}$, by fact 1 we have $\prod_U \mathfrak{N} \models \Diamond p_X \leftrightarrow p_{m(X)}$. Therefore $\prod_U \mathfrak{N}, s_1 \models p_{m(X)}$, and so $m(X) \in f(s_1)$ as desired.

Third, to show the 'zag' clause of bounded morphism, it suffices to show that $\{(s, f(s)) \mid s \in \prod_U \mathfrak{N}\}$ is a bisimulation. Now, both $\prod_U \mathfrak{N}$ and $\mathfrak{ue}\mathfrak{M}$ are modally saturated. By lemma 3.5, it suffices to show that for any $s \in \prod_U \mathfrak{N}$ and $u \in \mathfrak{ue}\mathfrak{M}$, $u = f(s)$ if and only if s and u are modally equivalent. Note that $\mathfrak{M} \models p_{V(\varphi)} \leftrightarrow \varphi$ by the definition of V, hence $\prod_U \mathfrak{N} \models p_{V(\varphi)} \leftrightarrow \varphi$ by fact 1. The rest of the proof is exactly the same as the one in Lemma 12.1.

Finally, we show that f is surjective. Let $u \subseteq \mathcal{P}(W)$ be an ultrafilter over W. We want to find $s \in \prod_U \mathfrak{N}$ such that $f(s) = u$. Let $\Theta = \{p_A \mid A \in u\}$. We show that Θ is finitely satisfiable in $\prod_U \mathfrak{N}$. Since u is an ultrafilter, $A_1 \cap \cdots \cap A_n \neq \emptyset$, and so $p_{A_1} \wedge \cdots \wedge p_{A_n}$ is satisfiable in $\mathfrak{M}$. By Fact 2, it is also satisfiable in $\prod_U \mathfrak{N}$. By ω−saturation, Θ is satisfiable in $\prod_U \mathfrak{N}$, hence there exists $s \in \prod_U \mathfrak{N}$ such that s satisfies Θ. We claim that $f(s) = u$, and the proof is again exactly the same as the one in Lemma 12.1. ■

12.3 Birkhoff's Theorem Revisited

Another way of addressing the 'purity of method' concern between algebra and model theory is to 'storm the capital city': Give model-theoretic proofs for the algebraic tools themselves, so that any algebraic proof involving them becomes 'model-theoretic' automatically. In the case of the Goldblatt-Thomason Theorem, if we can give a model-theoretic proof of Birkhoff's Theorem, then the entire proof would be model-theoretic. We give a sketch of how such a proof would go:

Proof. One direction of Birkhoff's theorem is straightforward: To show that an equational class of algebras is closed under taking products, subalgebras and homomorphic images, it suffices to show that the three algebraic operations preserve the satisfiability of equations.

For the other direction, suppose a class of algebras K is closed under taking products, subalgebras and homomorphic images, and we want to show that K is equationally definable. A natural candidate is the class of equations $\mathrm{Eq}(\mathsf{K})$ satisfied by K, since clearly $\mathsf{K} \subseteq \mathrm{Mod}(\mathrm{Eq}(\mathsf{K}))$. If we can show that $\mathrm{Mod}(\mathrm{Eq}(\mathsf{K})) \subseteq \mathsf{K}$, we will be able to conclude that $\mathsf{K} = \mathrm{Mod}(\mathrm{Eq}(\mathsf{K}))$. Now take

$\mathfrak{A} \vDash \mathrm{Eq}(\mathsf{K})$. To show that $\mathfrak{A} \in \mathsf{K}$, it suffices to show that $\mathfrak{A} \in \mathbf{HSP}\mathsf{K}$; that is: There exists $\{\mathfrak{A}_i\}_{i<\alpha} \subseteq \mathsf{K}$ and an algebra $\mathfrak{B}$, such that $\mathfrak{B}$ the subalgebra of the product $\mathbb{P} = \prod_{i<\alpha} \mathfrak{A}_i$, and that there is a surjective homomorphism from $\mathfrak{B}$ to $\mathfrak{A}$.

Let $\mathcal{L}$ be the language of $\mathrm{Eq}(\mathsf{K})$. Extend the language L to the language $\mathcal{L}[\mathfrak{A}]$ with constants c_a for each object a in $\mathfrak{A}$ (i.e., $\mathcal{L}[\mathfrak{A}] = \mathcal{L} \cup \{c_a\}_{a\in\mathfrak{A}}$), and let $T(\mathcal{L}[\mathfrak{A}])$ be the set of terms in $\mathcal{L}[\mathfrak{A}]$. Expand the algebra $\mathfrak{A}$ to a model $(\mathfrak{A}, a)_{a\in\mathfrak{A}}$, and define Σ to be the class of equations *falsified by* $(\mathfrak{A}, a)_{a\in\mathfrak{A}}$; that is, $\Sigma = \{t_1(\vec{x}) \approx t_2(\vec{x}) \mid t_1, t_2 \in T(\mathcal{L}[\mathfrak{A}]), (\mathfrak{A}, a)_{a\in\mathfrak{A}} \nvDash t_1(\vec{x}) \approx t_2(\vec{x})\}$. For every $\sigma \in \Sigma$, only finitely many new constants c_a $(a \in \mathfrak{A})$ occurs in σ. Let σ' be the result of replacing each new constant c_a with some variable x not already in σ, and it should be clear that $\mathfrak{A} \vDash \sigma'$ implies $(\mathfrak{A}, a)_{a\in\mathfrak{A}} \vDash \sigma$. We know that $(\mathfrak{A}, a)_{a\in\mathfrak{A}} \nvDash \sigma$; and so $\mathfrak{A} \nvDash \sigma'$. This implies that there exists $\mathfrak{A}_\sigma \in \mathsf{K}$ such that $\mathfrak{A}_\sigma \nvDash \sigma'$; Otherwise, $\sigma' \in \mathrm{Eq}(\mathsf{K})$, and so $\mathfrak{A} \vDash \sigma'$ after all. Moreover, we can define a $\mathcal{L}[\mathfrak{A}]$ expansion $(\mathfrak{A}_\sigma, c_a^{\mathfrak{A}_\sigma})_{a\in\mathfrak{A}}$ such that:

1. If c_a appears in σ, then we pick some element in $\mathfrak{A}_\sigma$ such that it is the relevant coordinate of a tuple falsifying σ' in $\mathfrak{A}_\sigma$. In other words: Since σ' is falsified in $\mathfrak{A}_\sigma$, we can always pick some tuple $\vec{\alpha}$ of elements of $\mathfrak{A}_\sigma$, such that $\mathfrak{A}_\sigma \nvDash \sigma'(\vec{\alpha})$. If the variable that replaces c_a is the ith variable in σ' then we define $c_a^{\mathfrak{A}_\sigma}$ as the i-th coordinate of the vector $\vec{\alpha}$.

2. If c_a does not occur in σ, then let $c_a^{\mathfrak{A}_\sigma}$ be an arbitrary element from $\mathfrak{A}_\sigma$.

Since $\mathfrak{A}_\sigma \nvDash \sigma'$, the definition of $(\mathfrak{A}_\sigma, c_a^{\mathfrak{A}_\sigma})_{a\in\mathfrak{A}}$ implies that $(\mathfrak{A}_\sigma, c_a^{\mathfrak{A}_\sigma})_{a\in\mathfrak{A}} \nvDash \sigma$.

Now given the set $\{\mathfrak{A}_\sigma : \sigma \in \Sigma\}$, we can construct the product $\mathbb{P} = \prod_{\sigma\in\Sigma} \mathfrak{A}_\sigma$, and $\mathbb{P} \in \mathsf{K}$ by the closure properties of K. (This is the $\mathbf{P}$ part of the construction.) Moreover, we can define a $\mathcal{L}[\mathfrak{A}]$-expansion of $\mathbb{P}$ as $(\mathbb{P}, c_a^{\mathbb{P}})_{a\in\mathfrak{A}}$, where $c_a^{\mathbb{P}} = (c_a^{\mathfrak{A}_{\sigma_1}}, c_a^{\mathfrak{A}_{\sigma_2}}, \dots)$; that is, $(\mathbb{P}, c_a^{\mathbb{P}})_{a\in\mathfrak{A}} = \prod_{\sigma\in\Sigma}(\mathfrak{A}_\sigma, c_a^{\mathfrak{A}_\sigma})_{a\in\mathfrak{A}}$. Since for every σ, $(\mathfrak{A}_\sigma, c_a^{\mathfrak{A}_\sigma})_{a\in\mathfrak{A}} \nvDash \sigma$, we have $(\mathbb{P}, c_a^{\mathbb{P}})_{a\in\mathfrak{A}} \nvDash \sigma$ for every $\sigma \in \Sigma$.

Next, we construct the algebra $\mathfrak{B}$. (This is the $\mathbf{S}$ part of the construction.) Let $\{c_a^{\mathbb{P}} \mid a \in \mathfrak{A}\}$ be the values of all the constants c_a in $(\mathbb{P}, c_a^{\mathbb{P}})_{a\in\mathfrak{A}}$, and let $\mathfrak{B}$ be the algebra generated by $\{c_a^{\mathbb{P}} \mid a \in \mathfrak{A}\}$. That is, $\mathfrak{B}$ is the smallest algebra containing $\{c_a^{\mathbb{P}} \mid a \in \mathfrak{A}\}$ that is closed under the old operations $\{f^{\mathbb{P}} \mid f \in \mathcal{L}\}$. By definition, $\mathfrak{B}$ is a subalgebra of $\mathbb{P}$, and so $\mathfrak{B} \in \mathsf{K}$ by the closure properties of K.

Finally, we show that $\mathfrak{A}$ is a homomorphic image of $\mathfrak{B}$. (This is the $\mathbf{H}$ part of the construction.) We define a map $h : \mathfrak{B} \to \mathfrak{A}$ such that

$$\text{For any } a \in \mathfrak{A},\ h(c_a^{\mathfrak{B}}) = a$$

$$\text{For any term } t \in T(\mathcal{L}) \text{ with variables } x_1, \ldots, x_n \text{ and any } c_{a_1}^{\mathfrak{B}}, \ldots, c_{a_n}^{\mathfrak{B}},$$
$$h(t^{\mathfrak{B}}(c_{a_1}^{\mathfrak{B}}, \ldots, c_{a_n}^{\mathfrak{B}})) = t^{\mathfrak{A}}(a_1, \ldots, a_n)$$

(We need some care as to how we think about terms in the algebra $\mathfrak{B}$, but we suppress details.)

The map h is guaranteed to be a surjective homomorphism from $\mathfrak{B}$ to $\mathfrak{A}$, and the only other thing we need to check is that h is well defined: That is, we need to check that

$$\text{If } t_1^{\mathfrak{B}}(c_{a_1}^{\mathfrak{B}}, \ldots, c_{a_n}^{\mathfrak{B}}) = t_2^{\mathfrak{B}}(c_{b_1}^{\mathfrak{B}}, \ldots, c_{b_m}^{\mathfrak{B}}), \text{ then } t_1^{\mathfrak{A}}(a_1, \ldots, a_n) = t_2^{\mathfrak{A}}(b_1, \ldots, b_m).$$

We prove this by contraposition: Suppose that $t_1^{\mathfrak{A}}(a_1, \ldots, a_n) \neq t_2^{\mathfrak{A}}(b_1, \ldots, b_m)$. Since $a_1, \ldots, a_n$, $b_1 \ldots b_m \in \mathfrak{A}$, this is equivalent to $\mathfrak{A} \not\models t_1(c_{a_1}, \ldots, c_{a_n}) \approx t_2(c_{b_1}, \ldots, c_{b_m})$. Hence $t_1(c_{a_1}, \ldots, c_{a_n}) \approx t_2(c_{b_1}, \ldots, c_{b_m}) \in \Sigma$. By our construction, $(\mathbb{P}, c_a^{\mathbb{P}})_{a \in \mathfrak{A}} \not\models t_1(c_{a_1}, \ldots, c_{a_n}) \approx t_2(c_{b_1}, \ldots, c_{b_m})$; that is, $t_1^{\mathbb{P}}(c_{a_1}^{\mathbb{P}}, \ldots, c_{a_n}^{\mathbb{P}}) \neq t_2^{\mathbb{P}}(c_{b_1}^{\mathbb{P}}, \ldots, c_{b_m}^{\mathbb{P}})$. Since $\mathfrak{B}$ is a restriction of $\mathbb{P}$ to the values of terms, we have $t_1^{\mathfrak{B}}(c_{a_1}^{\mathfrak{B}}, \ldots, c_{a_n}^{\mathbb{P}}) \neq t_2^{\mathfrak{B}}(c_{b_1}^{\mathfrak{B}}, \ldots, c_{b_m}^{\mathfrak{B}})$ as desired. ■

Further stratagems. Another example of an incursion into universal algebra is van Benthem's model-theoretic proof of Jónson's Lemma, a widely used algebaic result saying that, for varieties whose lattice of congruence relations are distributive, the subdirectly irreducible members belong to the class HSP_U of homomorphic images of subalgebras of ultraproducts of members of the class. The main point here is that such special algebras satisfy not just the algebraic equations valid in the class, but also the valid positive universal combinations of algebraic equations.

13 Infinitary Logic

There are also important logical systems other than FOL or SOL that have natural connections to modal logic. We discuss two such examples: infinitary logics in this section, and fixed pointed logics in the next section.

13.1 Infinitary language and semantics

$L_{\infty\omega}$ is a first-order language supplemented with arbitrary set conjunctions and disjunctions. To be precise, let τ be a set of constants, functions and relation symbols. The language $L_{\infty\omega}$ is built from τ, equality, Boolean connectives $\neg$, $\bigwedge$, $\bigvee$, quantifiers $\forall$ and $\exists$, and a set of variables:

- Terms and atomic formulas are defined as in first order logic.
- If ϕ is a formula, then so is $\neg\phi$.
- If X is a set of formula, then so are $\bigwedge_{\phi\in X}\phi$ and $\bigvee_{\phi\in X}\phi$.
- If ϕ is a formula, so are $\forall x\phi$ and $\exists x\phi$.

$L_{\omega\omega}$ is the fragment of $L_{\infty\omega}$ allowing only finite conjunctions and disjunctions (this is the usual first order language); $L_{\omega_1\omega}$ is the fragment that allows only countable conjunctions and disjunctions, etc. Given a natural number $2 \leq k < \omega$, the *k-variable fragment* $L^{(k)}_{\infty\omega}$ consists of those formulas of $L_{\infty\omega}$ which have been constructed using at most k constants and variables, free or bound.

The definition of an infinitary modal language $ML_{\infty\omega}$ is similar: Just take the standard modal language and add arbitrary set conjunctions and disjunctions.

We extend the usual definition of satisfiability by saying

$$\mathfrak{M} \models \bigwedge_{\phi\in X}\phi \text{ if and only if } \mathfrak{M} \models \phi \text{ for all } \phi \in X$$

and

$$\mathfrak{M} \models \bigvee_{\phi\in X}\phi \text{ if and only if } \mathfrak{M} \models \phi \text{ for some } \phi \in X$$

The notion of elementary equivalence can also be extended as

Definition 13.1. *1. Let $\mathfrak{M}$ and $\mathfrak{N}$ be L-structures. We write $\mathfrak{M} \equiv_{\infty\omega} \mathfrak{N}$ if*

$$\text{For all } L_{\infty\omega} \text{ sentences } \phi,\ \mathfrak{M} \models \phi \text{ iff } \mathfrak{N} \models \phi$$

2. Let $\vec{a} = (a_1, \ldots, a_n)$ be a sequence of n elements in $\mathfrak{M}$ and $\vec{b} = (b_1, \ldots, b_n)$ be a sequence of n elements in $\mathfrak{N}$. We write $(\mathfrak{M}, \vec{a}) \equiv_{\infty\omega} (\mathfrak{N}, \vec{b})$ if

$$\text{For all } L_{\infty\omega} \text{ formula } \phi(x_1, \ldots, x_n) \text{ with at most } n \text{ variables, } \mathfrak{M} \models \phi(\vec{a}) \text{ iff } \mathfrak{N} \models \phi(\vec{b})$$

In the following, we first prove two well-known theorems about $L_{\infty\omega}$, namely Karp's 'algebraic' characterization of $\equiv_{\infty\omega}$ in terms of potential isomorphism, and Scott's theorem that, for any $\mathfrak{M}$, there is a single $L_{\infty\omega}$ sentence that characterizes $\mathfrak{M}$ up to $\equiv_{\infty\omega}$. After going through these, we will explore the connection between infinitary logics and modal logic via interpolation and invariance.

13.2 Karp's Theorem

Definition 13.2. *Let $\mathfrak{M}$ and $\mathfrak{N}$ be L-structures. A **partial isomorphism** f from $\mathfrak{M}$ to $\mathfrak{N}$ is a partial function with domain contained in $\mathfrak{M}$ and range contained in $\mathfrak{N}$ that preserves function. That is, if $a_1, \ldots, a_n \in \mathrm{dom}(f)$ then for every n-ary predicate P of L,*

$$\langle a_1, \ldots, a_n \rangle \in P^{\mathfrak{M}} \text{ iff } \langle f(a_1), \ldots, f(a_n) \rangle \in P^{\mathfrak{N}}$$

Definition 13.3. *Let I be a non-empty set of partial isomorphisms from $\mathfrak{M}$ to $\mathfrak{N}$. We say that I is a **potential isomorphism** between $\mathfrak{M}$ and $\mathfrak{N}$, written $I : \mathfrak{M} \cong_p \mathfrak{N}$, if I satisfies the following back and forth property:*

1. *Forth: For every $f \in I$ and every $a \in \mathfrak{M}$, there is an $f' \in I$ such that $f \subseteq f'$ and a is in the domain of f'.*

2. *Back: For every $f \in I$ and every $b \in \mathfrak{N}$, there is an $f' \in I$ such that $f \subseteq f'$ and b is in the range of f'.*

Definition 13.4. *Let $\mathfrak{M}$ and $\mathfrak{N}$ be L-structures. The **Infinite Ehrenfeucht Game** $E(\mathfrak{M}, \mathfrak{N}, \infty)$ is played as follows:*

- *At stage n, Player I plays $a_n \in \mathfrak{M}$ or $b_n \in \mathfrak{N}$. In the first case Player II responds with $b_n \in \mathfrak{N}$ and in the second case Player II responds with $a_n \in \mathfrak{M}$.*

- *There is no bound on the number of moves; Player I and II alternate in making an ω-sequence of moves each.*

- *Player II wins if at each finite stage n of the play, the moves made so far (the map $a_n \mapsto b_n$) is a partial isomorphism.*

Theorem 13.5. *(Karp's Theorem). The following statements are equivalent:*

1. *$\mathfrak{M} \equiv_{\infty\omega} \mathfrak{N}$;*

2. $\mathfrak{M} \cong_p \mathfrak{N}$;

3. *There is a potential isomorphism I between $\mathfrak{M}$ and $\mathfrak{N}$ where every $f \in I$ has finite domain;*

4. *Player II has a winning strategy in $E(\mathfrak{M}, \mathfrak{N}, \infty)$.*

Proof. $(3 \Rightarrow 2)$: trivial.

$(2 \Rightarrow 1)$: Let I be a potential isomorphism between $\mathfrak{M}$ and $\mathfrak{N}$. We show that for all $n \in \mathfrak{N}$, all $\varphi(x_1, \ldots, x_n) \in L_{\infty\omega}$, all $f \in I$ and $\vec{a} = a_1, \ldots, a_n$ $(a_i \in \mathrm{dom}(f))$, we have

$$\mathfrak{M} \vDash \varphi(\vec{a}) \Longleftrightarrow \mathfrak{N} \vDash \varphi(f(\vec{a}))$$

We prove this by induction on formulas. This is clear for atomic formulas and the induction step is straightforward for $\neg, \bigwedge, \bigvee$.

Suppose $\varphi(\vec{x}) = \exists w \psi(\vec{x}, w)$ and $\mathfrak{M} \vDash \varphi(\vec{a})$. Then there exists $c \in \mathfrak{M}$ such that $\mathfrak{M} \vDash \psi(\vec{a}, c)$. By assumption, there exists $g \in I$ with $f \subseteq g$ and $c \in \mathrm{dom}(g)$. By the induction hypothesis, $\mathfrak{N} \vDash \psi(f(\vec{a}), f(c))$, and so $\mathfrak{N} \vDash \varphi(f(\vec{a}))$.

Conversely, suppose $\mathfrak{N} \vDash \varphi(f(\vec{a}))$. Then there exists $d \in \mathfrak{N}$ such that $\mathfrak{N} \vDash \psi(f(\vec{a}), d)$. By assumption, there exists $g \in I$ with $f \subseteq g$ and $c \in \mathrm{dom}(g)$ such that $g(c) = d$. By the induction hypothesis, $\mathfrak{M} \vDash \psi(\vec{a}, c)$, hence $\mathfrak{M} \vDash \varphi(\vec{a})$.

$(4 \Rightarrow 3)$: Let τ be a winning strategy for player II. Let I be the set of all maps f with finite domains such that $f(a_i) = b_i$, where $a_1, \ldots, a_n, b_1, \ldots b_n$ are the results of some game where at each stage player I has played either a_i or b_i and player II has responded using τ. Since τ is a winning strategy for player II, each such f is a partial isomorphism. Since player I can at any stage play any element from $\mathfrak{M}$ or $\mathfrak{N}$, I satisfies the definition of a potential isomorphism.

$(1 \Rightarrow 4)$: We need to prove a fact:

Fact: Suppose $(\mathfrak{M}, \vec{a}) \equiv_{\infty\omega} (\mathfrak{N}, \vec{b})$ and $c \in \mathfrak{M}$. Then there is $d \in \mathfrak{N}$ such that $(\mathfrak{M}, \vec{a}, c) \equiv_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$.

To prove this fact, suppose for contradiction that $(\mathfrak{M}, \vec{a}, c) \not\equiv_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$. Then for all $d \in \mathfrak{N}$, there is ϕ_d such that $\mathfrak{M} \vDash \phi_d(\vec{a}, c)$ and $\mathfrak{N} \nvDash \phi_d(\vec{b}, d)$. But then $\mathfrak{M} \vDash \exists w \bigwedge_{d \in \mathfrak{N}} \phi_d(\vec{a}, w)$ and $\mathfrak{N} \nvDash \exists w \bigwedge_{d \in \mathfrak{N}} \phi_d(\vec{b}, w)$, contradiction.

Given this fact, we can now describe Player II's strategy: She should play in such a way that after her n-th move a position $\{(a_1, b_1), \ldots, (a_n, b_n)\}$ is obtained, for which $(\mathfrak{M}, a_1, \ldots, a_n) \equiv_{\infty\omega} (\mathfrak{N}, b_1, \ldots, b_n)$. If she can do this (which is guaranteed by the assumption and the fact that we just proved), then the resulting map after her move will always be a partial isomorphism, and so she has a winning strategy. ■

The infinite Ehrenfeucht Game satisfies the conditions of the Gale-Stewart Theorem, which explains the above complementary powers of Players I and II.

13.3 Scott's Theorem

Now we turn to a result that extends and refines Karp's Theorem.

Definition 13.6. *Let $\mathfrak{M}$ be a model. For every finite sequence $\vec{a} = (a_1, \ldots, a_n)$ from $\mathfrak{M}$ and every ordinal γ, define the quantifier rank-γ formula $[\![\vec{a}]\!]^\gamma$ in the free variables $x_1, \ldots, x_n$, called the γ-**characteristic** of $\vec{a}$ in $\mathfrak{M}$, as follows:*

$$[\![\vec{a}]\!]^0_{\mathfrak{M}} = \bigwedge_{\psi \in X} \psi(\vec{x}), \text{ where } X = \{\psi : \mathfrak{M} \models \psi(\vec{a}) \text{ and } \psi \text{ is atomic or the negation of an atomic formula}\}.$$

$$[\![\vec{a}]\!]^\gamma_{\mathfrak{M}} = \bigwedge_{\beta < \gamma} [\![\vec{a}]\!]^\beta_{\mathfrak{M}}, \text{ where } \gamma \text{ is a limit ordinal.}$$

$$[\![\vec{a}]\!]^{\gamma+1}_{\mathfrak{M}} = \bigwedge_{m \in \mathfrak{M}} \exists x_{n+1} [\![\vec{a}, m]\!]^\gamma_{\mathfrak{M}} \wedge \forall x_{n+1} \bigvee_{m \in \mathfrak{M}} [\![\vec{a}, m]\!]^\gamma_{\mathfrak{M}}.$$

(Note that $[\![\vec{a}, m]\!]^\gamma_{\mathfrak{M}}$ has $n+1$ free variables: $x_1, \ldots, x_n$ and x_{n+1}.)

Definition 13.7. *Let $\mathfrak{M}$ and $\mathfrak{N}$ be L-models and $\vec{a} \in \mathfrak{M}^n$ and $\vec{b} \in \mathfrak{N}^n$ for arbitrary $n \in N$. For each ordinal γ, we define a relation $(\mathfrak{M}, \vec{a}) \equiv^\gamma_{\infty\omega} (\mathfrak{N}, \vec{b})$ recursively as follows:*

- $(\mathfrak{M}, \vec{a}) \equiv^0_{\infty\omega} (\mathfrak{N}, \vec{b})$ *if for all atomic formulas φ, $\mathfrak{M} \models \varphi(\vec{a}) \iff \mathfrak{N} \models \varphi(\vec{b})$.*
- *For all ordinals γ, $(\mathfrak{M}, \vec{a}) \equiv^{\gamma+1}_{\infty\omega} (\mathfrak{N}, \vec{b})$ if for all $c \in \mathfrak{M}$ there is $d \in \mathfrak{N}$ such that $(\mathfrak{M}, \vec{a}, c) \equiv^\gamma_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$ and for all $d \in \mathfrak{N}$ there is $c \in \mathfrak{M}$ such that $(\mathfrak{M}, \vec{a}, c) \equiv^\gamma_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$.*
- *For all limit ordinals γ, $(\mathfrak{M}, \vec{a}) \equiv^\gamma_{\infty\omega} (\mathfrak{N}, \vec{b})$ if $(\mathfrak{M}, \vec{a}) \equiv^\beta_{\infty\omega} (\mathfrak{N}, \vec{b})$ for all $\beta < \gamma$.*

Definition 13.8. *The **quantifier rank** $qr(\phi)$ of a $L_{\infty\omega}$ formula ϕ is defined as follows:*

- *If ϕ is atomic, $qr(\phi) = 0$.*

- *$qr(\neg\phi) = qr(\phi)$.*

- *$qr(\bigwedge_{\phi\in X} \phi) = qr(\bigvee_{\phi\in X} \phi) = \sup_{\phi\in X} qr(\phi)$.*

- *$qr(\forall x\phi) = qr(\exists x\phi) = qr(\phi) + 1$.*

Lemma 13.9. *$(\mathfrak{M}, \vec{a}) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b})$ if and only if*

$$\mathfrak{M} \vDash \varphi(\vec{a}) \iff \mathfrak{N} \vDash \varphi(\vec{b})$$

for all formulas $\varphi(\vec{x})$ of quantifier rank at most α.

Lemma 13.10. *Let $\mathfrak{M}, \mathfrak{N}$ be models such that $\vec{a}$ are in $\mathfrak{M}$ and $\vec{b}$ are in $\mathfrak{N}$. Then $\mathfrak{N} \vDash [\![\vec{a}]\!]^{\gamma}_{\mathfrak{M}}(\vec{b})$ if and only if $(\mathfrak{M}, \vec{a}) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b})$.*

Proof. We prove this by transfinite induction on γ. Because $(\mathfrak{M}, \vec{a}) \equiv^{0}_{\infty\omega} (\mathfrak{N}, \vec{b})$ if and only if they satisfy the same atomic formula, the definition of $[\![\vec{a}]\!]^{0}$ implies that $\mathfrak{N} \vDash [\![\vec{a}]\!]^{0}_{\mathfrak{M}}(\vec{b})$.

Suppose γ is a limit ordinal and the lemma is true for all $\beta < \gamma$. Then

$$\begin{aligned}
(\mathfrak{M}, \vec{a}) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b}) &\iff (\mathfrak{M}, \vec{a}) \equiv^{\beta}_{\infty\omega} (\mathfrak{N}, \vec{b}) \text{ for all } \beta < \gamma && (\text{def of } \equiv^{\gamma}_{\infty\omega}) \\
&\iff \mathfrak{N} \vDash [\![\vec{a}]\!]^{\beta}_{\mathfrak{M}}(\vec{b}) \text{ for all } \beta < \gamma && (\text{I.H}) \\
&\iff \mathfrak{N} \vDash [\![\vec{a}]\!]^{\gamma}_{\mathfrak{M}}(\vec{b}) && (\text{def of } [\![\vec{a}]\!]^{\gamma}_{\mathfrak{M}})
\end{aligned}$$

Suppose that the lemma is true for the formula γ. First, suppose that $\mathfrak{N} \vDash [\![\vec{a}]\!]^{\gamma+1}_{\mathfrak{M}}(\vec{b})$. Also, let $c \in \mathfrak{M}$. By assumption,

$$\mathfrak{N} \vDash \bigwedge_{m\in\mathfrak{M}} \exists x_{n+1} [\![\vec{a}, m]\!]^{\gamma}_{\mathfrak{M}}(\vec{b})$$

Hence there is a $d \in \mathfrak{N}$ such that $\mathfrak{N} \vDash [\![\vec{a}, c]\!]^{\gamma}_{\mathfrak{M}}(\vec{b}, d)$. By the induction hypothesis, $(\mathfrak{M}, \vec{a}, c) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$. Conversely, let $d \in \mathfrak{N}$. By assumption,

$$\mathfrak{N} \vDash \forall x_{n+1} \bigvee_{m\in\mathfrak{M}} [\![\vec{a}, m]\!]^{\gamma}_{\mathfrak{M}}(\vec{b})$$

And so there exists $c \in \mathfrak{M}$ such that $\mathfrak{N} \models [\![\vec{a}, c]\!]^{\gamma}_{\mathfrak{M}}(\vec{b}, d)$. Again by the induction hypothesis, $(\mathfrak{M}, \vec{a}, c) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$. Thus $(\mathfrak{M}, \vec{a}) \equiv^{\gamma+1}_{\infty\omega} (\mathfrak{N}, \vec{b})$.

Next, suppose that $(\mathfrak{M}, \vec{a}) \equiv^{\gamma+1}_{\infty\omega} (\mathfrak{N}, \vec{b})$. If $c \in \mathfrak{M}$, then there exists $d \in \mathfrak{N}$ such that $(\mathfrak{M}, \vec{a}, c) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$ and (by the induction hypothesis) $\mathfrak{N} \models [\![\vec{a}, c]\!]^{\gamma}_{\mathfrak{M}}(\vec{b}, d)$. Similarly, if $d \in \mathfrak{N}$, then there is $c \in \mathfrak{M}$ such that $\mathfrak{N} \models [\![\vec{a}, c]\!]^{\gamma}_{\mathfrak{M}}(\vec{b}, d)$. Thus $\mathfrak{N} \models [\![\vec{a}]\!]^{\gamma+1}_{\mathfrak{M}}(\vec{b})$ as desired. ■

Lemma 13.11. *For any infinite model $\mathfrak{M}$, there is an ordinal $\gamma < |\mathfrak{M}|^+$ such that if $\vec{a}, \vec{b} \in \mathfrak{M}^n$ and $(\mathfrak{M}, \vec{a}) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b})$, then $(\mathfrak{M}, \vec{a}) \equiv^{\beta}_{\infty\omega} (\mathfrak{N}, \vec{b})$ for all ordinal β. The least such γ is called the* ***Scott Rank*** *of $\mathfrak{M}$ and denoted as $sr(\mathfrak{M})$.*

Proof. For any ordinal σ, define $\Gamma_\sigma = \{(\vec{a}, \vec{b}) : \vec{a}, \vec{b} \in \mathfrak{M}^n$ for some $n = 0, 1, \ldots$ and $(\mathfrak{M}, \vec{a}) \not\equiv^{\sigma}_{\infty\omega} (\mathfrak{N}, \vec{b})\}$. Clearly, if $\sigma < \gamma$, $(\mathfrak{M}, \vec{a}) \not\equiv^{\sigma}_{\infty\omega} (\mathfrak{N}, \vec{b})$ implies $(\mathfrak{M}, \vec{a}) \not\equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b})$, and so $\Gamma_\sigma \subseteq \Gamma_\gamma$.

Claim 1: If $\Gamma_\sigma = \Gamma_{\sigma+1}$, then $\Gamma_\sigma = \Gamma_\gamma$ for all $\gamma > \sigma$.

We prove this by induction on γ. The base case is covered by the assumption $\Gamma_\sigma = \Gamma_{\sigma+1}$. If γ is a limit ordinal and the claim holds for all $\beta < \gamma$, then it also holds for γ. Now suppose the claim is true for $\gamma > \sigma$ and we want to show that it also holds for $\gamma + 1$. Since $\gamma + 1 > \sigma$, $\Gamma_\sigma \subseteq \Gamma_{\gamma+1}$. To show that $\Gamma_{\gamma+1} \subseteq \Gamma_\sigma$, it suffices to show that if $(\mathfrak{M}, \vec{a}) \equiv^{\sigma}_{\infty\omega} (\mathfrak{N}, \vec{b})$, then $(\mathfrak{M}, \vec{a}) \equiv^{\gamma+1}_{\infty\omega} (\mathfrak{N}, \vec{b})$. Suppose $(\mathfrak{M}, \vec{a}) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b})$ and $c \in \mathfrak{M}$. Since $\gamma > \sigma$, $(\mathfrak{M}, \vec{a}) \equiv^{\sigma+1}_{\infty\omega} (\mathfrak{N}, \vec{b})$, and so there exists $d \in \mathfrak{N}$ such that $(\mathfrak{M}, \vec{a}, c) \equiv^{\sigma}_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$. By the induction hypothesis, $(\mathfrak{M}, \vec{a}, c) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$. Similary, if $d \in \mathfrak{N}$, there exists $c \in \mathfrak{M}$ such that $(\mathfrak{M}, \vec{a}, c) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b}, d)$. Hence $(\mathfrak{M}, \vec{a}) \equiv^{\gamma+1}_{\infty\omega} (\mathfrak{N}, \vec{b})$ as desired.

Claim 2: There exists an ordinal $\gamma < |\mathfrak{M}|^+$ such that $\Gamma_\gamma = \Gamma_{\gamma+1}$.

Suppose not. Then for every $\gamma < |\mathfrak{M}|^+$, choose $(\vec{a}_\gamma, \vec{b}_\gamma) \in \Gamma_{\gamma+1}/\Gamma_\gamma$. The function $\gamma \mapsto (\vec{a}_\gamma, \vec{b}_\gamma)$ is one-to-one, since if $\sigma < \gamma$, $(\vec{a}_\sigma, \vec{b}_\sigma) \in \Gamma_{\sigma+1} \subseteq \Gamma_\gamma$ whereas $(\vec{a}_\gamma, \vec{b}_\gamma) \notin \Gamma_\gamma$. But this is a contradiction since there are only $|\mathfrak{M}|$ finite sequences from $\mathfrak{M}$. ■

Definition 13.12. *Let γ be the Scott rank of $\mathfrak{M}$. Define $\phi_{\mathfrak{M}}$ to be the sentence*

$$[\![\emptyset]\!]^{\gamma}_{\mathfrak{M}} \wedge \bigwedge_{\vec{a}} \forall \vec{x}([\![\vec{a}]\!]^{\gamma}_{\mathfrak{M}} \to [\![\vec{a}]\!]^{\gamma+1}_{\mathfrak{M}})$$

We call $\phi_{\mathfrak{M}}$ the ***Scott sentence*** *of $\mathfrak{M}$. (The subscript $\vec{a}$ means any finite sequence in $\mathfrak{M}$.)*

Theorem 13.13. ***(Scott's Theorem).*** *Let L be some infinitary language, $\mathfrak{M}$ be a L-structure and $\phi_{\mathfrak{M}}$ be the Scott sentence of $\mathfrak{M}$. Then for all L-structure $\mathfrak{N}$, $\mathfrak{N} \vDash \phi_{\mathfrak{M}}$ if and only if $\mathfrak{M} \equiv_{\infty\omega} \mathfrak{N}$.*

Proof. ($\Longleftarrow$): Let γ be the Scott rank of the model $\mathfrak{M}$. Then $\mathfrak{M} \vDash [\![\emptyset]\!]^{\gamma}_{\mathfrak{M}}$, since $[\![\emptyset]\!]^{\gamma}_{\mathfrak{M}}$ is a conjunction of sentences true in $\mathfrak{M}$. Moreover, suppose $\mathfrak{N} \vDash [\![\vec{a}]\!]^{\gamma}_{\mathfrak{M}}(\vec{b})$. By lemma 13.10, $(\mathfrak{M}, \vec{a}) \equiv^{\gamma}_{\infty\omega} (\mathfrak{N}, \vec{b})$. But since γ is the Scott rank of $\mathfrak{M}$, $(\mathfrak{M}, \vec{a}) \equiv^{\gamma+1}_{\infty\omega} (\mathfrak{N}, \vec{b})$ by lemma 13.11. Hence $\mathfrak{N} \vDash [\![\vec{a}]\!]^{\gamma+1}_{\mathfrak{M}}(\vec{b})$ and so $\mathfrak{N} \vDash \bigwedge_{\vec{a}} \forall \vec{x}([\![\vec{a}]\!]^{\gamma}_{\mathfrak{M}} \to [\![\vec{a}]\!]^{\gamma+1}_{\mathfrak{M}})$.

($\Longrightarrow$): suppose $\mathfrak{N} \vDash \phi_{\mathfrak{M}}$. We will show that $\mathfrak{M} \cong_q \mathfrak{N}$. Let I be the set of functions with finite domains from $\mathfrak{M}$ to $\mathfrak{N}$ such that if $\vec{a}$ is in the domain of f and $f(\vec{a}) = \vec{b}$, then $\mathfrak{N} \vDash [\![\vec{a}]\!]^{\gamma}(\vec{b})$. Since $\mathfrak{N} \vDash \phi_{\mathfrak{M}}$, $\mathfrak{N} \vDash [\![\emptyset]\!]^{\gamma}$. Hence the empty map is in I and I is nonempty.

Suppose $f \in I$ and $f(\vec{a}) = \vec{b}$, and let $c \in \mathfrak{M}$. Since $\mathfrak{N} \vDash \phi_{\mathfrak{M}} \wedge [\![\vec{a}]\!]^{\gamma}(\vec{b})$, $\mathfrak{N} \vDash [\![\vec{a}]\!]^{\gamma+1}(\vec{b})$. By the definition of $[\![\vec{a}]\!]^{\gamma+1}$, $\mathfrak{N} \vDash \exists x_{n+1}[\![\vec{a}, c]\!]^{\gamma}(\vec{b})$; that is, there exists d such that $\mathfrak{N} \vDash [\![\vec{a}, c]\!]^{\gamma}(\vec{b}, d)$ and so we can extend f by sending c to d.

On the other hand suppose we are given $d \in \mathfrak{N}$. We know that $\mathfrak{N} \vDash [\![\vec{a}]\!]^{\gamma+1}(\vec{b})$, which implies $\mathfrak{N} \vDash \forall x_{n+1} \bigvee_{m \in \mathfrak{M}} [\![\vec{a}, m]\!]^{\gamma}$. Then there exists $c \in \mathfrak{M}$ such that $\mathfrak{N} \vDash [\![\vec{a}, c]\!]^{\gamma}(\vec{b}, d)$, and we can extend f be sending c to d.

Thus I is a partial isomorphism and $\mathfrak{M} \equiv_p \mathfrak{N}$. By Karps' theorem, $\mathfrak{M} \equiv_{\infty\omega} \mathfrak{N}$. ∎

We conclude this subsection with another version of Scott's theorem that will be used in proving the interpolation theorem for $L_{\infty\omega}$:

Theorem 13.14. ***(Scott's Theorem: Version 2):*** *Let $\mathfrak{M}$ be a L-structure. For every ordinal γ there is an infinitary sentence $\phi^{\gamma}_{\mathfrak{M}}$ such that for all L-structures $\mathfrak{N}$, $\mathfrak{N} \vDash \phi^{\gamma}_{\mathfrak{M}}$ if and only if $\mathfrak{M} \equiv^{\gamma}_{\infty\omega} \mathfrak{N}$. Moreover, the collection of all such sentences $\phi^{\gamma}_{\mathfrak{M}}$ for all $L-$structures $\mathfrak{M}$ forms a set rather than a proper class.*

Proof. We show that the sentence $\phi^{\gamma}_{\mathfrak{M}} = [\![\emptyset]\!]^{\gamma}_{\mathfrak{M}}$ satisfies the condition.

($\Longleftarrow$): Since $\mathfrak{M} \vDash [\![\emptyset]\!]^{\gamma}_{\mathfrak{M}}$, $[\![\emptyset]\!]^{\gamma}_{\mathfrak{M}}$ has quantifier rank γ and $\mathfrak{M} \equiv^{\gamma}_{\infty\omega} \mathfrak{N}$, $\mathfrak{N} \vDash [\![\emptyset]\!]^{\gamma}_{\mathfrak{M}}$.

($\Longrightarrow$): Suppose $\mathfrak{N} \vDash [\![\emptyset]\!]^{\gamma}_{\mathfrak{M}}$, and we want to show that $\mathfrak{M} \equiv^{\gamma}_{\infty\omega} \mathfrak{N}$. We prove this by induction on γ. The case where $\gamma = 0$ is trivial. If γ is a limit ordinal, then for all $\beta < \gamma$, $\mathfrak{N} \vDash [\![\emptyset]\!]^{\beta}_{\mathfrak{M}}$. By the induction hypothesis, $\mathfrak{M} \equiv^{\beta}_{\infty\omega} \mathfrak{N}$ for all $\beta < \gamma$, and so $\mathfrak{M} \equiv^{\gamma}_{\infty\omega} \mathfrak{N}$. If $\gamma = \sigma + 1$, then $\mathfrak{N} \vDash \bigwedge_{c \in \mathfrak{M}} \exists x_{n+1} [\![c]\!]^{\sigma}_{\mathfrak{M}}$ and $\mathfrak{N} \vDash \forall x_{n+1} \bigvee_{c \in \mathfrak{M}} [\![c]\!]^{\sigma}_{\mathfrak{M}}$. That is, for every $c \in \mathfrak{M}$, there exists $d \in \mathfrak{N}$ such that $\mathfrak{N} \vDash [\![c]\!]^{\sigma}_{\mathfrak{M}}(d)$ and for every $d \in \mathfrak{M}$, there exists $c \in \mathfrak{M}$ such that $\mathfrak{N} \vDash [\![c]\!]^{\sigma}_{\mathfrak{M}}(d)$. Lemma 13.10 implies that , for every $c \in \mathfrak{M}$, there exists $d \in \mathfrak{N}$ such that $(M, c) \equiv^{\sigma}_{\mathfrak{M}} (\mathfrak{N}, d)$ and for for every $d \in \mathfrak{M}$, there exists $c \in \mathfrak{M}$ such that $(M, c) \equiv^{\sigma}_{\mathfrak{M}} (\mathfrak{N}, d)$. By definition, $\mathfrak{M} \equiv^{\gamma}_{\mathfrak{M}} \mathfrak{N}$ as desired. ■

13.4 Undefinability and Restoring Interpolation

While all our results so far suggest strong expressive power for $L_{\infty\omega}$, it is also important to realize that it does have strong limitations. Here is a classical result by Lopez-Escobar from the 1970s.

Theorem 13.15. *Well-foundedness of a binary ordering is not definable by a sentence of $L_{\infty\omega}$.*

While this may seem a negative result, it actually does play a role as a substitute for the first-order Compactness Theorem, a model existence result that obviously fails for $L_{\infty\omega}$. Here is a generalized Lindström Theorem illustrating this role.

Theorem 13.16. *$L_{\infty\omega}$ is the strongest extension of first-order logic satisfying (a) invariance for potential isomoprhism, and (b) undefinability of well-ordering.*

Next, whether a property 'holds' or 'fails' for a logical system may depend in delicate ways on how it is formulated. To see this, we discuss one more crucial feature of $L_{\infty\omega}$, its behavior with respect to the basic first-order property of interpolation.

First, recall the standard formulation of the interpolation theorem:

Definition 13.17. *A language L has the* ***standard interpolation property*** *if for all $\varphi, \psi \in L$, the following are equivalent:*

1. *φ entails ψ.*
2. *There is a sentence $\theta \in L$ in the vocabulary $L_{\varphi} \cap L_{\psi}$ such that $\varphi \vDash \theta$ and $\theta \vDash \psi$.*

First order logic $L_{\omega\omega}$ has the standard interpolation property, so does $L_{\omega_1\omega}$.

Theorem 13.18. *$L_{\infty\omega}$ does not have the standard interpolation property.*

We do not give an example of such a failure here, but it is easy to find one, using the existence of two formulas in $L_{\infty\omega}$ with disjoint vocabulary defining a well-order of type ω and one of type ω_1, respectively: for details, see [2].

Before jumping to the conclusion that $L_{\infty\omega}$ really lacks interpolation, however, we need to ask: Is the formulation of the property we obtained from first-order logic the correct one to generalize? The same property may be formulated in multiple ways, all equivalent in FOL; but it is entirely possible that some of the formulations will hold for other systems while others will not.

For instance, the standard formulation of the interpolation property in Theorem 84 is peculiar, in the sense that intuitively, 2 seems to say something stronger than 1. It allows us to make a jump from one model to another, provided that these are suitably related. In [2], Barwise and van Benthem calls this general notion 'entailment along a relation', and ordinary consequence is a special case of it, namely, the identity relation of staying in one model. Here is a formal version:

Definition 13.19. *Let r be a binary relation on L-structures. We say that φ **entails** ψ **along** R, written $\varphi \vDash_R \psi$, iff for all L-structures $\mathfrak{M}$ and $\mathfrak{N}$, if $\mathfrak{M} R \mathfrak{N}$ and $\mathfrak{M} \vDash \varphi$ then $\mathfrak{N} \vDash \psi$. In particular, φ is **preserved under** R if $\varphi \vDash_R \varphi$.*

In $L_{\infty\omega}$, although the standard formulation of interpolation theorem fails, we can formulate a alternative version using the general notion of 'entailment along a relation':

Theorem 13.20. *(Interpolation Theorem for $L_{\infty\omega}$). Given any two formulas $\varphi, \psi \in L_{\infty\omega}$, the following statements are equivalent:*

1. *φ entails ψ along potential isomorphism.*
2. *There is an ordinal α such that φ entails ψ along $\equiv^{\alpha}_{\infty\omega}$.*
3. *There is a sentence θ of $L_{\infty\omega}$ such that $\varphi \vDash \theta$ and $\theta \vDash \psi$.*

Proof. $(3 \to 1)$: This follows immediately from the left-to-right direction of Karp's Theorem.

$(1 \to 2)$: Let α be the quantifier rank of φ, and let $\mathfrak{M} \vDash \varphi$ and $\mathfrak{M} \equiv^{\alpha}_{\infty\omega} \mathfrak{N}$. Since $\varphi \in L_{\infty\omega}$, $\mathfrak{N} \vDash \varphi$. It then follows from (1) that $\mathfrak{N} \vDash \psi$.

$(2 \to 3)$: Let α be the ordinal in (2). By Scott's theorem, for every model $\mathfrak{M}$ such that $\mathfrak{M} \vDash \varphi$, there exists a sentence $\sigma^{\alpha}_{\mathfrak{M}}$ that characterizes $\mathfrak{M}$ up to $\equiv^{\alpha}_{\infty\omega}$, and $\chi = \{\sigma^{\alpha}_{\mathfrak{M}} \mid \mathfrak{M} \vDash \varphi\}$ is a set. Hence $\theta = \bigvee_{\sigma \in \chi} \sigma$ is a sentence of $L_{\infty\omega}$. To see that $\varphi \vDash \theta$, let $\mathfrak{M}$ be any model of φ. Then $\mathfrak{M} \vDash \sigma^{\alpha}_{\mathfrak{M}}$, and this is one of the disjunction of θ, so $\mathfrak{M} \vDash \theta$. To see that $\theta \vDash \psi$, let $\mathfrak{N} \vDash \theta$, and so there exists $\mathfrak{M} \vDash \varphi$ such that $\mathfrak{N} \vDash \sigma^{\alpha}_{\mathfrak{M}}$. By Scott's theorem, $\mathfrak{M} \equiv^{\alpha}_{\infty\omega} \mathfrak{N}$. By (2), $\mathfrak{N} \vDash \psi$ as desired. ■

In [2], Barwise and van Benthem prove a still more generalized version of Theorem 13.20, and they provide a number of applications.

13.5 Infinitary Modal Logic

One reason why infinitary logics are very natural from the view point of modal logic is that they provide a new perspective on the modal invariance theorem. In this final subsection, we show that a modal invariance theorem (similar to the one for basic modal logic) holds for $L_{\infty\omega}$, and that it follows directly from a special version of a suitable *interpolation theorem* for $ML_{\infty\omega}$. The main reason is that the definitions and proofs in the preceding subsection specialize to this modal fragment of full infinitary logic.

To illustrate this, consider the the formulation of the interpolation theorem as applied to modal language. Suppose that 2 is the case. Then not only is ψ a consequence of φ, but something stronger is also the case, namely that ψ is a 'consequence of φ along a bisimulation':

> Let $\mathfrak{M}$ be a model of φ. Let $\mathfrak{N}$ be another model such that $\mathfrak{M}$ and $\mathfrak{N}$ are bisimilar as far as the shared language $L_\varphi \cap L_\psi$ goes. Then, if we have an interpolant in the shared language such that $\varphi \vDash \alpha$ and $\alpha \vDash \psi$, we also know that $\mathfrak{M}$ is a model of α. Now, since there is a bisimulation between $\mathfrak{M}$ and $\mathfrak{N}$ with regard to the language of α, $\mathfrak{N}$ must be a model of α. It follows that $\mathfrak{N}$ is a model of ψ.

Theorem 13.21. *(Interpolation Theorem for $ML_{\infty\omega}$). The following are equivalent for all $\varphi, \psi \in L_{\infty\omega}$ both with one free variable x:*

1. *φ entails ψ along bisimulation.*
2. *There exists some formulas χ in $ML_{\infty\omega}$ with $\varphi \vDash \chi$ and $\chi \vDash \psi$.*

Proof. See [2] for details of how the earlier interpolation argument also works when using bisimulation in the place of potential isomorphism. ∎

Corollary 13.22. *(Modal Invariance Theorem for $L_{\infty\omega}$)*

A formula $\varphi(x)$ of $L_{\infty\omega}$ is invariant for bisimulation iff it is equivalent to a formula of $ML_{\infty\omega}$.

Proof. Suppose that φ is invariant for bisimulation, that is, φ entails φ along bisimulation. By Theorem 13.21, φ is equivalent to a $ML_{\infty\omega}$-formula—namely its $ML_{\infty\omega}$ interpolant. ∎

There are many further interesting aspects to infinitary modal logic when we study more concrete issues. For instance, by an argument just as for $L_{\infty\omega}$, there are modal Scott-formulas defining pointed models up to bisimulation, whose form is reminiscent of 'nabla formulas' in co-algebraic logic that enumerate modal types occurring in the model. But also, such formulas can have very special formats for specific classes of models. We merely cite one fact, from [18], Section 5:

Theorem 13.23. *Each finite pointed model $\mathfrak{M}, s$ has a formula in propositional dynamic logic defining that model up to bisimulation.*

Interestingly, propositional dynamic logic is a small and widely used fragment of $ML_{\infty\omega}$.

14 Fixed Point Logics

Infinitary modal languages are good for defining induction and recursion while staying bisimulation-invariant. However, a perhaps more elegant direct approach using only finitary syntax is that of fixed-point logics such as LFP(FO) or in the modal case, the modal mu-calculus. We start with some basics about LFP(FO), the fixed-point extension of first-order logic.

14.1 Language and semantics of LFP(FO)

LFP(FO) is the extension of FOL with operators for smallest and greatest fixed points:

Definition 14.1. *The language of* LFP(FO) *extends the usual formation rules for first-order syntax with an operator defining smallest fixed points (and one defining greatest fixed points):*

$$[(\mu P, \overline{x}).\phi(\overset{+}{P}, \overline{Q}, \overline{x})](\bar{t})$$
$$[(\nu P, \overline{x}).\phi(\overset{+}{P}, \overline{Q}, \overline{x})](\bar{t})$$

where P may occur only positively in $\phi(\overset{+}{P}, \overline{Q}, \overline{x})$; $\overline{x}$ is a tuple of variables and $\bar{t}$ is a tuple of terms, both with the same length as that of the arity of P.

Definition 14.2. *The semantics of the new formulas are:*

$$\mathfrak{M} \models [(\mu P, \overline{x}).\phi(\overset{+}{P}, \overline{Q}, \overline{x})](\overline{a}) \textit{ iff } \overline{a} \textit{ is in the least fixed point of } F_\phi$$

$$\mathfrak{M} \models [(\nu P, \overline{x}).\phi(\overset{+}{P}, \overline{Q}, \overline{x})](\overline{a}) \textit{ iff } \overline{a} \textit{ is in the greatest fixed point of } F_\phi$$

where F_ϕ *is the monotone operator induced by* $\phi(\overset{+}{P}, \overline{Q}, \overline{x})$ *defined by*

$$F_\phi(X) = \{\overline{a} \mid \mathfrak{M} \models \phi(\overset{+}{P}, \overline{Q}, \overline{x})[X, \overline{a}]\}$$

To prove that this definition works as stated we need to show that, if P occurs only positively in $\phi(\overset{+}{P}, \overline{Q}, \overline{x})$, then F_ϕ is in fact a monotone operator. The proof is a simple induction, in which one must also take care of the fixed-point operators themselves.

By the well-known Tarski-Knaster Theorem, every monotone operator has a least fixed point and a greatest fixed point on any complete lattice with smallest and greatest element - and thus, the semantics of the new LFP(FO)-formulas is well-defined.

First-order fixed-point logic is a very natural and elegant system that has been used to define many kinds of induction and recursion, from computation to games, cf. [22]. One striking difference with infinitary logic $L_{\infty\omega}$ is that LFP(FO) can define well-foundedness of binary orderings in one simple smallest fixed-point formula.

As for general properties, like infinitary first-order logic, LFP(FO) has the Karp Property:

LFP(FO)-formulas are invariant for potential isomorphism.

Proof. This result can be proved by a straightforward induction on formulas. ■

Another important property is this. LFP(FO) satisfies the downward Löwenheim-Skolem Theorem in the following strong form:

Each model $\mathfrak{M}$ for a formula ϕ with a countable submodel $\mathfrak{M}'$
has a countable submodel containing $\mathfrak{M}'$ where ϕ is still true.

Proof. The proof for this result is by encoding the standard stagewise ordinal approximation process for all fixed-points occurring in ϕ in a first-order theory, while adding first-order ordering properties for the ordinal ordering as well as its well-foundedness. The resulting theory has a countable model of the right sort by the ordinary Löwenheim-Skolem Theorem plus the downward persistence of well-foundedness, and this is the required model for ϕ. ■

Incidentally, the construction in this proof (due to Joerg Flum) can be used to show the central role of well-foundedness once more: validity for LFP(FO) is mutually faithfully interpretable with validity in first-order logic with an one additional constant proposition for well-foundedness.

Apart from these two properties, little is known about the general model theory of LFP(FO), and in particular, no Lindström Theorem has been found so far for this system.

14.2 Minimization, Intersection Property and PIA Syntax

Before returning to modal logic, we consider one major reformulation of LFP(FO) in terms of minimizing (or maximizing) extensions of predicates. As pointed out by van Benthem in [19], from which the material in this subsection is taken, this is often useful in applications to modal correspondence theory, or even other areas such as non-monotonic logic.

We are after the existence of minimal predicates satisfying a certain first-order description, and the crucial semantic property ensuring this is as follows.

Definition 14.3. *A first-order formula $\phi(P, \overline{Q})$ has the **intersection property** for the predicate letter P if, in any model $\mathfrak{M}$, whenever $\mathfrak{M}$, holds for all predicates in some family $\{P_i \mid i \in I\}$, it also holds for their intersection. That is, if $\mathfrak{M}, P_i \vDash \phi(P, \overline{Q})$ for all $i \in I$, then $\mathfrak{M}, \cap P_i \vDash \phi(P, \overline{Q})$. Here $\overline{Q}$ is a tuple of predicate letters in the vocabulary, and $(\mathfrak{M}, P_i)$ is the same model as $\mathfrak{M}$ except that the interpretation of the predicate letter P is P_i.*

Now, here is a syntactic form which guarantees that the intersection property will hold.

Definition 14.4. *A first-order formula is a PIA-condition (short for 'positive antecedent implies atom') if it has the following syntactic form*

$$\forall \overline{x}(\varphi(\overset{+}{P}, \overline{Q}, \overline{x}) \to P\overline{x})$$

with P occurring only positively in $\varphi(\overset{+}{P}, \overline{Q}, \overline{x})$. Here $\overline{Q}$ is again a tuple of letters and $\overline{x}$ a tuple of individual variables.

[19] shows that a first order formula has the intersection property if and only if it is definable by means of PIA formulas, so PIA is the syntax needed for the argument for the intersection property.

We proceed to give some details of this result, since it provides some further and not so well-known understanding of the model theory of fixed-point logics.

Lemma 14.5. *All PIA-conditions $\phi(P, \overline{Q})$ have the intersection property.*

Proof. Suppose $\mathfrak{M}, P_i \vDash \phi(P, \overline{Q})$ for all $i \in I$, where $\phi(P, \overline{Q}) = \forall \overline{x}(\varphi(\overset{+}{P}, \overline{Q}, \overline{x}) \to P\overline{x})$. Now let $\mathfrak{M}, \cap P_i, \overline{d} \vDash \varphi(\overset{+}{P}, \overline{Q}, \overline{x})$ where $\overline{d}$ is a tuple of objects. By the positive occurrence of P in $\varphi(\overset{+}{P}, \overline{Q}, \overline{x})$, we have $\mathfrak{M}, P_i, \overline{d} \vDash \varphi(\overset{+}{P}, \overline{Q}, \overline{x})$ for all $i \in I$. But then $\mathfrak{M}, P_i, \overline{d} \vDash P\overline{x}$ for all $i \in I$, and hence $\mathfrak{M}, \cap P_i, \overline{d} \vDash P\overline{x}$ as desired. ■

Theorem 14.6. *The following are equivalent for all first-order formulas $\phi(P, \overline{Q})$:*

1. *$\phi(P, \overline{Q})$ has the intersection property with respect to predicate P.*

2. *$\phi(P, \overline{Q})$ is definable by means of a PIA formula with respect to P.*

Proof. $(2 \to 1)$: This follows directly from Lemma 14.5.

$(1 \to 2)$: Assume condition (1). For notational simplicity, we take P to be a unary predicate. Now define the following set of syntactic consequences of ϕ:

$$\text{PIA-Cons}(\phi) = \{\psi \mid \psi \text{ is PIA with respect to } P \text{ and } \phi \vDash \psi\}.$$

Our goal is to show that

$$\text{PIA-Cons}(\phi) \vDash \phi$$

If we can show this, then by the compactness theorem, ϕ is implied by some finite conjunction of its own PIA-consequences with respect to P, and hence it is equivalent to this conjunction. Condition (2) then follows because any such conjunction is equivalent to a single PIA formula.

We now prove PIA-Cons$(\phi) \vDash \phi$. Let $\mathfrak{M}$ be any model of the language $L(P, \overline{Q})$ satisfying PIA-Cons(ϕ). First, we dispose of a special case. If $\mathfrak{M} \vDash \forall x Px$, then $\phi(P, \overline{Q})$ holds automatically in $\mathfrak{M}$. To see this, note that by the intersection property of ϕ, $\mathfrak{M}, \cap\emptyset \vDash \phi(P, \overline{Q})$. But $\cap\emptyset$ is the whole domain of $\mathfrak{M}$ and by assumption everything in $\mathfrak{M}$ satisfies P. Hence $\mathfrak{M}$ and $(\mathfrak{M}, \cap\emptyset)$ are the same model, and so $\mathfrak{M} \vDash \phi(P, \overline{Q})$.

For the remainder of the proof, we assume that $\mathfrak{M} \nvDash \forall x Px$. To fix notation, let $L(\overline{Q})$ be the first-order language with the base predicates $\overline{Q}$ only, and similarly $L(P, \overline{Q})$ is the language with predicates P and $\overline{Q}$. A series of model constructions gives us the following result:

Setup Lemma. There is an elementary extension $\mathfrak{M}^*$ of $\mathfrak{M}$ and, for each $d \in \mathfrak{M}^*$ that does not satisfy P, a model $\mathfrak{N}_d$ and a map f_d from $\mathfrak{M}^*$ to $\mathfrak{N}_d$ such that

1. ϕ is true in $\mathfrak{N}_d$.
2. $f_d(d)$ does not satisfy P in $\mathfrak{N}_d$.
3. f_d is an $L(\overline{Q})$-isomorphism and a P-(weak) homomorphism from $\mathfrak{M}^*$ onto $\mathfrak{N}_d$.

The setup lemma has a long proof and we refer to [19] for details. For now, note the consequences of this lemma: Each model N_d satisfies ϕ. Moreover, let $P_d = \{f_d^{-1}(e) \mid e \in P^{\mathfrak{N}_d}\}$, which can be thought of as the interpretation of the predicate P in $\mathfrak{M}^*$ by coping from the interpretation of P in $\mathfrak{N}_d$ via the map f_d. The P-homomorphism condition ensures that P_d contains $P^{\mathfrak{M}^*}$. Also, it should be clear that f_d is an $L(P, \overline{Q})$-isomorphism between the model $(\mathfrak{M}^*, P_d)$ and $\mathfrak{N}_d$. Finally, since $f_d(d)$ does not satisfy P in $\mathfrak{N}_d$, $d \notin P_d$.

Now we use the given intersection property of ϕ. We know that $\mathfrak{M}^*, P_d \vDash \phi$ for every d that does not satisfy P in $\mathfrak{M}^*$. Hence $\mathfrak{M}^*, \cap P_d \vDash \phi$. But by the preceding observations, $\cap P_d = P^{\mathfrak{M}^*}$: Clearly $P^{\mathfrak{M}^*} \subseteq \cap P_d$. Moreover, if $a \in \cap P_d$ but $a \notin P^{\mathfrak{M}^*}$, then there exists a set P_a such that $a \notin P_a$, which is impossible since $\cap P_d \subseteq P_a$. Therefore, $\mathfrak{M}^* \vDash \phi$. But then also $\mathfrak{M} \vDash \phi$, since $\mathfrak{M}^*$ is an elementary extension of the original model $\mathfrak{M}$ for PIA-Cons(ϕ) ■

It should be noted that what we have established here a preservation result for first-order logic. There is a natural further question whether the above theorem also holds for the full language of LFP(FO), for which the above Compactness-based proof does not work. This is an open problem, as are much of the usual model theoretic first-order preservation theorems.

Next, there is a fairly transparent connection between PIA syntax and fixed point logics: What the PIA formula $\forall \overline{x}(\varphi(\overset{+}{P}, \overline{Q}, \overline{x}) \to P\overline{x})$ says about P is that P is a pre-fixed point of the operator induced by $\varphi(\overset{+}{P}, \overline{Q}, \overline{x})$. But if we can define pre-fixed points in PIA syntax, we can also talk about μ-formulas and smallest fixed point. Such an intuitive connection can be made precise as follows:

Definition 14.7. ***(Language of*** MIN(FO)***)***

The language of first-order logic with predicate minimization MIN(FO) *and the set of extended PIA-conditions are defined by simultaneous induction as the least sets such that:*

1. *Every first-order formula is a* MIN(FO)*-formula.*

2. *Every PIA-condition is an extended PIA-condition.*

3. *Whenever there is an extended PIA-condition $\phi(P, \overline{Q})$ and a tuple of terms $\bar{t}$ with the same length as the arity of P, then* $[(\text{MIN } P).\phi(P, \overline{Q})](\bar{t})$ *is a* MIN(FO)*-formula.*

4. *Formulas of the form $\forall\overline{x}(\psi(\overset{+}{P}, \overline{Q}, \overline{x}) \to P\overline{x})$, where $\psi(\overset{+}{P}, \overline{Q}, \overline{x})$ is a* MIN(FO)*-formula and P occurs only positively in $\psi(\overset{+}{P}, \overline{Q}, \overline{x})$, is an extended PIA-condition. (A relational symbol R occurs only positively in* MIN(FO)*-formula $[(\text{MIN } P).\phi](\bar{t})$, for $R \neq P$, iff it occurs only positively in ϕ.)*

Theorem 14.8. MIN(FO) *and* LFP(FO) *have the same expressive power, that is, for every* MIN(FO)*-formula there is a* LFP(FO)*-formula with the same models and vice versa.*

Proof. We merely give a sketch.

From LFP(FO) to MIN(FO). We prove this by induction. The base case is straightforward. Now suppose $\psi(\overset{+}{P}, \overline{Q}, \overline{x})$ is a LFP(FO) formula that has a MIN(FO)-equivalent $\gamma(\overset{+}{P}, \overline{Q}, \overline{x})$. By the proof of the Tarski-Knaster Theorem, the smallest fixed-point for an operator is also the smallest pre-fixed point for that operator. Hence the smallest fixed-point formula can be represented in MIN(FO) as follows:

$$[(\mu P, \overline{x}).\psi(\overset{+}{P}, \overline{Q}, \overline{x})](\bar{t}) = [(\text{MIN } P).\forall\overline{x}(\gamma(\overset{+}{P}, \overline{Q}, \overline{x}) \to P\overline{x})](\bar{t})$$

From MIN(FO) to LFP(FO). Again the proof is by induction. Let $\forall\overline{x}(\gamma(\overset{+}{P}, \overline{Q}, \overline{x}) \to P\overline{x})$ be an extended PIA-condition, and suppose the MIN(FO)-formula $\gamma(\overset{+}{P}, \overline{Q}, \overline{x})$ has a LFP(FO)-equivalent $\psi(\overset{+}{P}, \overline{Q}, \overline{x})$. But then $[(\text{MIN } P).\forall\overline{x}(\gamma(\overset{+}{P}, \overline{Q}, \overline{x}) \to P\overline{x})](\bar{t})$ can be represented in LFP(FO) as $[(\mu P, \overline{x}).\psi(\overset{+}{P}, \overline{Q}, \overline{x})](\bar{t})$. ■

14.3 Frame Correspondence in LFP(FO)

Returning now to our modal themes, fixed-point logics are natural to work with from the point of view of modal frame correspondence, as certain correspondence properties that are not definable in FO are definable in fixed-point logics.

For instance, consider Löb's Axiom

$$\Box(\Box p \to p) \to \Box p$$

which says that the accessibility relation is transitive and upward well-founded. Its corresponding defining formula is actually in LFP(FO): Transitivity is first order, and converse well foundedness can be defined even in μ-calculus by $\mu p.\Box p$. So the corresponding property of Löb's axiom is relatively simple in fixed-point logic.

Can we find frame correspondents in LFP(FO) as systematically as first-order frame correspondents? Recall how we found first-order correspondents using Sahlqvist Theorem: Look at the antecedent of the modal formula, and try to find a "minimal valuation" where the antecedent is true. If the minimal valuation can be expressed by a first order formula, we plug it into (the standard translation of) the consequent and get a first-order correspondent. The two key ideas here were the minimal valuation itself and the definability of the minimal valuation in first order logic. These two things can come apart, however, as we shall see in the case of Löb's Axiom.

There is a minimal way of making the antecedent of Löb's Axiom true: Suppose

$$\Box(\Box p_i \to p_i)$$

is true for a family of proposition letters p_i ($i \in I$). Then the following is also true:

$$\Box(\Box \textstyle\bigwedge_{i\in I} p_i \to \bigwedge_{i\in I} p_i)$$

What this means is that, if we take all valuations for p that makes the antecedent of Löb's Axiom true and intersect them, the resulting (minimal) valuation will still make the antecedent true.

Of course, what we have established here is the Intersection Property of the preceding subsection. And what we also recognize, looking at the syntactic form of Löb's Axiom, is that, under the standard first-order translation, it exhibits PIA syntax, which ensures that it has the intersection property and hence a minimal valuation. By our earlier results, this ensures automatic definability in LFP(FO), via a simple transformation.

Theorem 14.9. *Modal axioms with PIA antecedents and syntacticially positive consequents all have their corresponding frame conditions definable in LFP(FO).*

Proof. We merely provide a sketch. Let $\alpha \to \beta$ be a modal formula such that the standard translation $ST_x(\alpha)$ is equivalent to a PIA-condition $\forall y(\phi(\overset{+}{P}, x, y) \to Py)$ with respect to predicate P, and $ST_x(\beta) = \psi(\overset{+}{P}, x)$ is equivalent to a formula in which P has only positive occurrences. Since PIA-conditions have intersection property, there is a minimal value of P satisfying $\forall y(\phi(\overset{+}{P}, x, y) \to Py)$, which is the intersection of all the values of P satisfying $\forall y(\phi(\overset{+}{P}, x, y) \to Py)$. This minimal value can be represented by a MIN(FO)-formula $[(\text{MIN } P).\forall y(\phi(\overset{+}{P}, x, y) \to Py)](t)$ or, equivalently, by a LFP(FO)-formula $[(\mu P, y).\phi(\overset{+}{P}, x, y)](t)$. By substituting this LFP(FO)-formula for each of the predicate letter P in the consequent $ST_x(\beta)$, we arrive at a LFP(FO)-formula γ.

The proof that γ is indeed the LFP(FO)-correspondent of $\alpha \to \beta$ is basically the same as the proof that Sahlqvist's algorithm works. ■

To illustrate the algorithm, consider Löb's Axiom $\Box(\Box p \to p) \to \Box p$. The standard translation of the antecedent is equivalent to a PIA-condition $\forall y(Rxy \wedge \forall z(Ryz \to Pz) \to Py)$, and the standard translation of the consequence is a formula $\forall w(Rxw \to Pw)$ in which P only occurs positively. The minimal value for the PIA condition is $[(\mu P, y).(Rxy \wedge \forall z(Ryz \to Pz))](t)$, and by substituting this for Pw in $\forall w(Rxw \to Pw)$, we obtain a LFP(FO)-formula $\forall w(Rxw \to [(\mu P, y).(Rxy \wedge \forall z(Ryz \to Pz))](w))$. Interpreted semantically, this says precisely that R is transitive at x and that there is no infinite sequence of R-successors starting from x.

Finally, there is still a level of frame correspondence beyond first-order logic and LFP(FO). We are now also in a position to make sense of our earlier remark that "the McKinsey Axiom is less well behaved than Löb's Axiom" in Section 6.2. We saw that the correspondent of Löb's Axiom ends up in LFP(FO), since the antecedent is of PIA form. In contrast, the modal antecedent $\Box\Diamond p$ of McKinsey Axiom has a non-PIA quantifier pattern: $\forall x(Rxy \to \exists y(Ryz \wedge ...)$. This is significant, and it turns out that McKinsey Axiom is *not* LFP(FO)-definable:

Theorem 14.10. *The modal McKinsey Axiom $\Box\Diamond p \to \Diamond\Box p$ has no LFP(FO)-definable frame correspondent.*

Proof. Recall the uncountable frame $\mathfrak{F}$ defined in the proof of Theorem 34, where we showed that $\mathfrak{F} \models \Box\Diamond p \to \Diamond\Box p$, while this formula failed in some suitably chosen countable elementary subframe. But LFP(FO) satisfies the downward Skolem-Löwenheim Theorem, as we noted earlier (see [9]).

Hence, if $\Box\Diamond p \to \Diamond\Box p$ is definable by a LFP(FO)-formula, then there is a countable LFP(FO) frame $\mathfrak{F}^-$ (as in the proof of Theorem 6.2) such that $\mathfrak{F}^- \models \Box\Diamond p \to \Diamond\Box p$. But we also verified that $\mathfrak{F}^- \not\models \Box\Diamond p \to \Diamond\Box p$, contradiction. Hence the McKinsey Axiom is not definable in LFP(FO). ■

14.4 Modal μ-Calculus

There is a well-known modal counterpart to the first-order fixed-point logic, viz. the modal μ-calculus that extends ML with operators for smallest and greatest fixed-points. This system has the usual features of a modal logic, being bisimulation-invariant, and even definable. Here we just make a few remarks on correspondence aspects.

First, we can extend the above analysis to the language of the μ-calculus rather than that of basic modal logic. In this case, too, minimal valuations, and suitable extended PIA syntax make sense, and general Sahlqvist-type results can be proved. We can allow arbitrary fixed-point operators in the positive consequents, and specially constrained occurrences of both smallest and greatest fixed-point operators in antecedents. For the somewhat complex details of the syntax of these "Sahlqvist μ-formulas", we refer to [23].

Theorem 14.11. *All μ-Sahlqvist axioms have LFP(FO)-definable frame correspondents.*

But there is also another line of thought here, given the new expressive power. For instance, the correspondence result about L/'ob's Axiom can now be stated as the following validity in the object language of the μ-calculus:

Lób's Logic is axiomatized alternatively by $\mu p.\Box p \ \wedge \ (\Box p \to \Box\Box p$

The reason is that the single μ-formula $\mu p.\Box p$, when used as an axiom, defines well-foundedness of a relation. The general topic arising here is that of "internalizing" semantic correspondence arguments into suitable proof calculi (cf. [14]), some times even extended modal systems. One reason is that modal fixed-point logics allow for non-trivial and sometimes surprising formal deductions, witness the discussion of connections between the μ-calculus and Lób's provability logic in [20].

Finally, some model-theoretic results discussed for the basic modal logic in these notes are open problems in this realm. For instance, it is unknown whether there is a Goldblatt-Thomason

Theorem for the μ-calculus, where, to keep harmony on both sides, we might restrict attention to the modally μ-definable LFP(FO) frame classes closed under suitable frame-building operations.

15 Conclusion

These lecture notes have provided a brief tour of the classical model theory of modal logic, with a few modern additions toward infinitary and fixed-point logics. Our emphasis was on ideas and general methods, and not all these topics have been discussed in detail, while open problems abound (for instance, about the intersection of infintary and fixed-point extensions of modal logic).

Even at the level of general perspectives, we have left out some major trends in current research that would fit naturally with what we have discussed. We mention a few:

- Algebraic perspectives on all of the above,
- Proof-theoretic perspectives on all of the above,
- Category-theoretic perspectives on all of the above (especially, in the field of "co-algebra"),
- Automata-theoretic perspectives on all of the above (especially relevant to our current understanding of the modal μ-calculus and monadic second-order logic),
- Extensions to modal neighborhood semantics (where point-to-point relations of our graph-like frames are replaced by point-to-set relations as in hyper-graphs),
- Extensions to the "Guarded Fragment", [1], which lifts the first-order base of modal logic to a much larger fragment that remains decidable even when fixed-point operators are added.

References

[1] H. Andréka, J. van Benthem & I. Németi, 1998. "Modal Logics and Bounded Fragments of Predicate Logic", *Journal of Philosophical Logic*, vol. 27:3, 217274.

[2] Barwise, J. and van Benthem, J., 1999, "Interpolation, Preservation, and Pebble Games", *The Journal of Symbolic Logic*, Vol. 64, No. 2, 881-903.

[3] Bell, J. L. and Slomson, A, 1969, *Models and Ultraproducts*, North-Holland Publishing Company, Amsterdam.

[4] Blackburn, P., De Rijke, M and Venema, Y., 2001. *Modal Logic*, Cambridge University Press, Cambridge.

[5] Burris, S. and Sankappanavar, H.P., 1981, *A Course in Universal Algebra*, Springer-Verlag.

[6] Chang, C. and Keisler, J., 1990, *Model Theory*, third edition. Studies in Logic and the Foundations of Mathematics, 73. North-Holland Publishing Co., Amsterdam.

[7] Cirera, R. F., 2011, *An Introduction to Model Theory*, `http://www-ma2.upc.edu/farre/index_fitxers/Model-Theory.pdf`

[8] Doets, K., 1996, *Basic Model Theory*. CSLI Publications.

[9] Grädel, E., 1999, "The Decidability of Guarded Fixed Point Logic", `http://www.illc.uva.nl/j50/contribs/graedel/graedel.pdf`

[10] Keisler, J., "The Ultraproduct Construction". `http://www.math.wisc.edu/~keisler/ultraproducts-web-final.pdf`

[11] Marker, D., 2013, *Lectures on Infinitary Model Theory*, `homepages.math.uic.edu/~marker/math512-F13/inf.pdf`

[12] Mierzewski, C., "Ultraproducts in Model Theory", Department of Philosophy, Stanford University.

[13] van Benthem, J., 1977, *Modal Correspondence Theory*, Dissertation, Mathematical Institute, University of Amsterdam.

[14] van Benthem, J., 1979, "Syntactic Aspects of Modal Incompleteness Theorems", *Theoria*, vol. 45:2, 63-77.

[15] van Benthem, J., 1984, "Modal Correspondence Theory", in D. Gabbay and F. Guenthner, eds., *Handbook of Philosophical Logic*, vol. II, 167-247.

[16] van Benthem, J., 1985, *Modal Logic and Classical Logic*, Bibliopolis, Napoli.

[17] van Benthem, J., 1997, "Modality, Bisimulation and Interpolation in Infinitary Logic", *Annals of Pure and Applied Logic*, Vol.96, 29-41.

[18] van Benthem, J., 1998, "Dynamic Odds and Ends", Tech report, Institute for Logic, Language and Computation, University of Amsterdam.

[19] van Benthem, J., 2005, "Minimal Predicates, Fixed Points, and Definability", *The Journal of Symbolic Logic*, Vol.70, No. 3.

[20] van Benthem, J., 2006, "Modal Frame Correspondences and Fixed-Points", *Studia Logica*. vol. 83:1, 133155.

[21] van Benthem, J., 2010, *Modal Logic for Open Minds*, CSLI Publications, Stanford CA.

[22] van Benthem, J., 2014, *Logic in Games*, The MIT Press, Cambridge MA.

[23] van Benthem, J., Bezhanishvili, N., Hodkinson, I., 2012, "Sahlqvist Correspondence for Modal Mu-Calculus", *Studia Logica* Vol.100: 3160

[24] van Benthem, J., Bezhanishvili, N., Holliday, W., 2015, "A Bimodal Perspective on Possibility Semantics", *Logic and Computation*, to appear.