INDICES

Monographs in Philosophical Logic and Formal Linguistics

3

JOHAN VAN BENTHEM

MODAL LOGIC AND CLASSICAL LOGIC

BIBLIOPOLIS

INDICES

Monographs in Philosophical Logic and Formal Linguistics

III

JOHAN VAN BENTHEM

MODAL LOGIC AND CLASSICAL LOGIC

BIBLIOPOLIS

ISBN 88-7088-113-X

Photoset by «Studio in/out»
Milano
Printed in Italy by «Grafitalia s.r.l.»
Cercola, November 1985

CONTENTS

PREFACE

The present work is a rewritten version of my dissertation *Modal Correspondence Theory* ([5]) and a supplementary report called *Modal Logic as Second-Order Logic* ([7]). I would like to repeat the acknowledgments made in the former.

The assistance of my "promotor" M. H. Löb and my "co-referent" S. K. Thomason has been invaluable, for their constructive criticism. I am also very grateful to Mrs. M. G. Eichhorn-Pigge who typed both dissertation and report. In fact, the entire staff of the "Vakgroep Logika en Grondslagen van de Wiskunde" at the University of Amsterdam has been a constant source of encouragement, advice and often active help. I mention W. J. Blok, H. C. Doets, D. H. J. de Jongh (who helped me very generously at a critical stage in writing the dissertation) and A. S. Troelstra. As for people active in the field of modal logic, I have learned much from the work of K. Fine, R. I. Goldblatt, H. Sahlqvist, K. Segerberg and S. K. Thomason.

It should be added that this book owes its existence to the patient labours of J. Perzanowski of the Jagiellonian University at Krakow. Owing to circumstances beyond his and my control, the road from manuscript (completed in 1978) to published work has been a long and tortuous one, involving travels from Holland via Poland to Italy. It was F. Guenthner who supervised the final stage of that journey, leading to the present publication.

Still, my greatest debt is owed to my parents, to who I dedicate this work:

Bram van Benthem
Anne Eggermont.

INTRODUCTION

The research which has led to this publication started in 1972. It was inspired by a remark in Krabbe [42], a discussion of K. Segerberg's paper [69]. The latter paper is devoted to proving completeness theorems for modal logics, of which the following result is a simple example. A modal formula belongs to the normal modal logic axiomatized by $Lp \rightarrow LLp$ (where "L" stands for "necessarily") if and only if it holds on all Kripke frames whose alternative relation is transitive. (A modal formula is said to hold on a Kripke frame if it is true there under every valuation on that frame.) Now, Krabbe observed that, on any Kripke frame, $Lp \rightarrow LLp$ holds if and only if the alternative relation of that frame is transitive. In other words, he gave a *direct* correspondence between the modal formula $Lp \rightarrow LLp$ and the first order sentence $\forall x \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))$ defining transitivity.

Since the operator L behaves like a restricted universal quantifier in Kripke semantics ("for all R-successors of the world in which evaluation takes place"), there is a syntactic similarity between the above two formulas, in that both contain a nested sequence of two such quantifiers. This simple observation started off a search for a syntactic translation procedure taking those modal formulas that express first-order relational properties to the first-order formulas describing these properties. A paper by F. B. Fitch ([26]), constructing such equivalents for modal formulas of the form $M^k p \rightarrow M_1 \dots M_l p$ (where $k \geqslant 0, l \geqslant 1$, "$M$" stands for "possibly", "M^k" stands for "M ... (k times) ... M" and $M_i = L$ or $M_i = M$ $(1 \leqslant i \leqslant l)$), served as a further stimulus. Unfortunately, in 1973 it turned out that the results obtained in this way had been given already by H. Sahlqvist (cf. [66]). The method by which they were obtained remains of some independent interest, however; whence it is presented in chapter 9 below.

When formulating syntactic criteria for a modal formula to define a first-order property (a necessary condition for the above search), it becomes important to have a method for showing that certain modal formulas do *not* define such properties. Such a method, based upon an application of the Löwenheim-Skolem theorem, was found, and published in [1] to show that the well-known formula $LMp \rightarrow MLp$ does not define a first-order property of the alternative relation. In the report [10] a summary was given of the results obtained at that stage (spring 1974) together with some applications to, e. g., tense logic. That summer, a combined application of both methods mentioned above yielded a complete syntactic answer to the question which *modal reduction principles* define first-order properties. Modal reduction principles — the name is due to Fitch — are modal formulas of the form $M_1 \ldots M_k p \rightarrow M_{k+1} \ldots M_{k+l} p$ (where $k \geqslant 0$, $l \geqslant 0$ and $M_i = L$ or M $(1 \leqslant i \leqslant k + 1)$). Note that many well-known modal axioms have this form: $Lp \rightarrow p$, $Lp \rightarrow LLp$, $MLp \rightarrow p$, etc. This result was published in [8].

An unpublished paper ([2]) about a model-theoretic treatment of literary interpretation, and the discovery of an error in it by S. K. Thomason, led to a semantic characterization of the modal formulas defining first-order properties as those being preserved under ultrapowers. The original involved proof of this fact was presented in a simplified form in [6], using a lemma due to R. I. Goldblatt (cf. chapter 8 below).

In 1975, it was decided to use the three papers [1], [6] and [8] — all of which had appeared in the *Journal of Symbolic Logic* — as a basis for a dissertation about the connection between modal formulas and first-order relational properties. They were accompanied by a general survey of this subject, which grew into a book ([5]) containing several new topics, such as the preservation results of chapter 15 below. Throughout this book, modal formulas are presented as a special kind of second-order formulas. This point of view was motivated in the introduction by referring to S. K. Thomason's proof (in [77]) that the second-order notion of semantic consequence is reducible to that of modal logic. (For more details, cf. chapter 1 below.) It was emphasized even more strongly in the subsequent report [7] containing the "parerga et paralipomena" which could not be incorporated into [5] any more. In fact, a large part of [7] is devoted to higher-order logic alone. (Cf. part IV of this book.)

Most results referred to up to now have been obtained by means of *model-theoretic* methods. *Algebraic* methods (although used so very ably by R. A. Bull (cf. [16]) and D. C. Makinson (cf. [55]) seemed less promising. But the dissertations of R. I. Goldblatt ([31]) and W. J. Blok ([14]) showed that an algebraic point of view may be quite useful in matters of definability. (E. g., Birkhoff's characterization of equational varieties becomes available.) Some applications of Goldblatt's technique will be found in chapter 16 below. (One particular such application — a characterization of the so-called *canonical* modal logics — has appeared as [3].) Of course, the difference between model-theoretic and algebraic methods is more apparent than real, because, after all, (universal) algebra may be viewed as that part of model theory in which predicates are disregarded.

This research may be said to be concerned with *definability theory*. The modal language is studied as a means of defining properties of the alternative relation, and a comparison is made with other languages serving the same purpose. E. g. the following questions will be treated below.

(1) When does a given modal formula define a first-order property (of the alternative relation)?

(2) When can a given first-order property be defined by means of a modal formula (or a set of such formulas)?

(3) Which classes of frames are definable at all by means of modal formulas?

To all three questions, model-theoretic answers exist (cf. chapters 8, 14 and 16, respectively). It has been hoped that a purely syntactic answer to (1) and (2) would be possible as well; but several results in chapters 7 and 9 make it seem plausible that (at least) the class of modal formulas defining a first-order relational property is not even arithmetically definable.

The main subject in modal logic up to now has been a different one, however, viz. *completeness theory*. The classical reference on this subject is Segerberg [67]. Here this subject will be mentioned in chapter 6 only, in order to discuss its connection with definability theory. This connection turns out to be rather involved, and a lot of open questions remain in this area. Completeness theory recently received new impetus from the discovery of so-called *incomplete* modal logics by K. Fine ([22]) and S. K. Thomason ([75]). It is treated thoroughly from an algebraic point of view in a series of publications by W. J. Blok.

This book is organized in the following manner. Part I provides a short survey of the standard part of propositional modal logic, including the notions and results needed for our definability theory. In chapter 7, which starts off part II, two notions of first-order definability for modal formulas will be defined, one "local" and one "global". Surprisingly, these turn out to be different. The method of ultraproducts yields a semantic characterization for these formulas in chapter 8. Their syntactic study is taken up in chapter 9; but, again, semantic methods are needed in chapter 10 to delimit the range of these syntactic methods. Two more special subjects, first-order definability modulo some condition on the alternative relation, and first-order definability for formulas of modal predicate logic, are treated in chapters 11 and 12. An idea from chapter 9 inspired the general concept of a "preservation class" of modal formulas, studied in chapter 13, which, in a sense, connects definability theory with completeness theory. In part III, the emphasis shifts to first-order formulas. To begin with, in chapter 14, first-order formulas defined by modal formulas are characterized semantically. Some syntactic results about such formulas are found in chapter 15. The more general question which classes of frames are modally definable is treated in chapter 16. Finally, in the last part of the book, the notions and results obtained at this stage are applied to higher-order languages of increasing strength, viz. universal second-order logic, second-order logic in general and the theory of finite types.

The techniques and notions of the present work could be applied to various branches of intensional logic; e. g., intuitionistic logic, relevance logic, or even quantum logic. In fact there is a publication about first-order definability of a modal language when the alternative relation is taken to be ternary, or n-ary for arbitrary n (cf. [38]). Moreover, P. Rodenburg has recently written a dissertation on intuitionistic correspondence theory. Still, possible generalizations to higher-order logic would seem to be most interesting, because they provide a connection with the mainstream of classical logic.

It remains to be mentioned that a compact, but accessible survey of the present theory — containing various additional results — has appeared as a chapter called "Correspondence Theory" in volume 2 of the *Handbook of Philosophical Logic* (edited by D. Gabbay and F. Guenthner, Reidel, Dordrecht, 1984).

NOTATION AND TERMINOLOGY

The usual set-theoretic and model-theoretic notation will be used freely. Moreover, the abbreviations $\Rightarrow$ (*if ... then ...*), $\Leftrightarrow$ (*if and only if*), & (*and*), $\sim$ (*not*), $\forall$ (*for every*) and $\exists$ (*there exists*) will occur occasionally to facilitate the informal exposition. Standard results from classical logic are presupposed (cf. Chang & Keisler [17], Enderton [20] or Shoenfield [72]). Moreover, routine proofs are omitted, most of them being simple inductions with respect to the complexity of (modal) formulas.

The modal language to be studied is that of modal propositional logic. (Modal predicate logic will be treated in chapter 12 only.) It contains an infinite supply of proposition letters $p_1, p_2, ..., p, q,...$ The Boolean operators are $\neg$ (*not*), $\wedge$ (*and*), $\vee$ (*or*), $\rightarrow$ (*if ... then...*) and $\leftrightarrow$ (*if and only if*). Moreover, there are two propositional constants $\bot$ (the so-called *falsum*, standing for a fixed contradiction) and $\top$ (the so-called *verum*, standing for a fixed tautology). When stating definitions or proving propositions, it is often convenient to use a minimal set of primitives. $\neg$ and $\wedge$ are selected for this purpose here. But, there are some cases where another choice of primitives (e.g., $\neg$ and $\rightarrow$ or $\bot$ and $\rightarrow$) has been preferred as yielding more elegant formulations. Two unary modal operators L (*necessarily*) and M (*possibly*) are added to the Boolean ones. Either one suffices as a primitive concept, L being definable as $\neg M\neg$ and M as $\neg L\neg$. In fact, L is taken to be primitive here, so $\neg, \wedge, L$ is the standard set of primitives for the modal language. *Modal formulas* are obtained in the usual way, and small Greek letters φ, ψ, χ,... (possibly with subscripts) will be used to denote formulas. Modal formulas wich are constructed from $\bot$ and $\top$ without using proposition letters will be called *closed formulas*. Finally, Greek capitals Γ, Δ, Σ,... (possibly with subscripts) will be used for sets of formulas.

The first-order language referred to in the Introduction contains one, binary predicate constant R as well as identity $=$. Its individual variables are x_1, x_2,..., x, y, z, u, v, w. Its Boolean operators are as above, and the quantifiers are $\forall$ (for every) and $\exists$ (there exists). This language is called L_0. Formulas of L_0 will be denoted by α, β, γ,...(possibly with subscripts).

L_1 is the first-order language with one binary predicate constant R and unary predicate constants P_1, P_2,..., P, Q,... corresponding to the proposition letters of the modal language. When these unary predicate constants are regarded as predicate *variables*, the second-order language L_2 is obtained. Note that L_2 has only one first-order predicate constant, viz. R. In the final part IV, the full language of first-order and second-order predicate logic will be needed, but the four languages introduced here suffice for the earlier parts.

In the semantics for the above languages, the following model-theoretic notation will be used. Structures will be denoted by bold letters. E.g., for the modal language we need *frames* consisting of a non-empty domain (the set of "worlds") together with a binary relation on that domain (the "alternative relation" or "relation of accessibility"). Notation: $\mathbf{F}$ $(= \langle W, R\rangle)$, $\mathbf{F}_1$ $(= \langle W_1, R_1\rangle)$, ... Note that frames are also structures for L_0 and L_2. In addition, auxiliary structures called *models* are used for the modal language. Models consist of a frame $\mathbf{F}$ together with a *valuation* V on $\mathbf{F}$ mapping proposition letters onto subsets of W. ($V(p)$ is to be thought of as the set of worlds where the state of affairs p obtains.) Notation: $\mathbf{M}$ $(= \langle \mathbf{F}, V\rangle$ or, alternatively, $\langle W, R, V\rangle)$, $\mathbf{M}_1$ $(= \langle \mathbf{F}_1, V_1\rangle)$, ... Note that models may be regarded as structures for L_1 in an obvious way. Finally, there are *general frames* consisting of a frame $\mathbf{F}$ $(= \langle W, R\rangle)$ together with a set $\mathbf{W}$ of subsets of W which is closed under intersections, relative complements (with respect to W) and the set-theoretic operation l defined by $l(X) = \{w \in W \mid \forall v \in W\ (Rwv \Rightarrow v \in X)\}$ for all $X \subseteq W$. Notation: $\langle \mathbf{F}, \mathbf{W}\rangle (= \langle W, R, \mathbf{W}\rangle)$, $\langle \mathbf{F}_1, \mathbf{W}_1\rangle$ $(= \langle W_1, R_1, \mathbf{W}_1\rangle)$, ... General frames are also structures for L_2 in the extended sense in which Henkin's *general models* (cf.[33]) are structures for the full second-order language. ($\mathbf{W}$ restricts the range of the predicate quantifiers.) In this perspective, frames correspond to standard models for L_2 and may thus be identified with "full" general frames in which $\mathbf{W}$ is the power set of W.

The truth definition in chapter 2 yields the central notion $\mathbf{M} \models \varphi[w]$ (φ *holds/is true at* w *in* $\mathbf{M}$), where $\mathbf{M}$ is a model $\langle W, R, V\rangle$, $w \in W$ and φ is a modal formula. Several derived notions: $\mathbf{M} \models \varphi$, and, for a set Σ

of modal formulas, $\mathbf{M} \models \Sigma\ [w]$ and $\mathbf{M} \models \Sigma$ are then defined in an obvious way. A useful notation is $Th_{mod}(\mathbf{M})$ for $\{\varphi \mid \mathbf{M} \models \varphi\}$. Next, $\mathbf{F} \models \varphi[w]$ (*φ holds at w in* $\mathbf{F}$) is defined for any frame $\mathbf{F}$ $(= \langle W, R \rangle)$ and $w \in W$. Again, $\mathbf{F} \models \varphi$, $\mathbf{F} \models \Sigma[w]$, $\mathbf{F} \models \Sigma$ and $Th_{mod}(\mathbf{F})$ are defined in terms of this. In accordance with practice in standard logic, the same sign $\models$ is also used for the purpose of denoting *universal validity* and *semantic consequence.* $\models \varphi$ means "φ is universally valid" and $\Sigma \models \varphi$ means "φ follows semantically from Σ". This last notion $\models$ is split up into several different ones (e.g., $\models_f$ and $\models_m$) in chapter 2. Since it will not be used very often, this will cause no confusion.

Several semantic notations involving frames will be introduced in chapter 2, such as $\mathbf{F}_1 \underset{\rightarrow}{\subseteq} \mathbf{F}_2$ ($\mathbf{F}_1$ is a *generated subframe* of $\mathbf{F}_2$), $\Sigma\,\{\mathbf{F}_i \mid i \in I\}$ (the *disjoint union* of $\{\mathbf{F}_i \mid i \in I\}$) and $ue(\mathbf{F})$ (the *ultrafilter extension* of $\mathbf{F}$). Going back to that chapter will often help a confused reader. For algebraic notation the reader is referred to chapter 4.

The syntactic theory of modal logic (cf. chapters 5 and 6) employs only a few general notations. The *modal logic axiomatized by* a set Σ of modal formulas is denoted by $ML(\Sigma)$. Modal axiomatic theories receive their common names, e.g., K, T, $S4$, $S5$. *Derivability* within a theory is denoted by $\vdash_K$, $\vdash_T$, etc. C is the class of *complete* modal logics, GC that of the modal logics which are *generally complete*, while CAN is the class of the *canonical* modal logics.

Finally, the following notation will be used frequently in the discussion of the correspondence between modal formulas and L_0-formulas (cf. chapter 3). $E(\varphi, \alpha)$ denotes *local equivalence* between a modal formula φ and an L_0-formula α. Likewise, $\bar{E}(\varphi, \alpha)$ denotes *global equivalence*. $M1$ is the set of modal formulas with a local equivalent in L_0, $\bar{M}1$ the set of those having a global equivalent in L_0. Dually, $P1$ consists of those L_0-formulas which have an equivalent in $M1$, and $\bar{P}1$ of the L_0-formulas with an equivalent in $\bar{M}1$.

The remaining notations are more ad hoc, and will be introduced in the course of the exposition.

PART I

A Short Survey Of Propositional Modal Logic

CHAPTER I

HISTORICAL BACKGROUND

Modal logic was born in philosophy. The logic of necessity, possibility and contingence was already studied in Antiquity (e. g., by Aristotle) and the Middle Ages (cf. Bochenski [15]). Its modern revival is due to C. I. Lewis (cf. [49]) in the twenties of this century. Lewis was dissatisfied with the *material* implication of Russell & Whitehead's *Principia Mathematica*. According to him, "A implies B" as used in ordinary discourse means more than just "It is not the case that (A and not B)". He therefore added the so-called *strict* implication to the *Principia* system, which may be paraphrased as "It is *impossible* that (A and not B)", or — equivalently — "It is necessarily true that not (A and not B)". Here we are only interested in the notions of possibility and necessity arising from these considerations. Let the unary modal operators M ("possibly") and L ("necessarily") be added to the language of propositional logic, as explained above. Modern modal logic started with a search for the "correct" axioms describing the logical behaviour of L and M. Certain principles were relatively unproblematic, like the interdefinability of L and M ($Lp \leftrightarrow \neg M \neg p$, $Mp \leftrightarrow \neg L \neg p$), the modalized form of Modus Ponens ($L(p \rightarrow q) \rightarrow (Lp \rightarrow Lq)$), or the implication $Lp \rightarrow p$. But it soon turned out that not one, but several axiomatic theories ("modal logics") were possible, some of which have become well-known, such as $S4$ and $S5$ (with characteristic axioms — in addition to the above three — $Lp \rightarrow LLp$ and $MLp \rightarrow Lp$, respectively). From then on, modal logic consisted in the study of hosts of axiomatic theories, a phenomenon which has continued to this day.

In the thirties, it was tried to give an algebraic semantics for these theories, but, although some interesting results were reached (cf. the somewhat later publication [58], the field remained rather chaotic. Unlike in standard logic, where the Tarski semantics led to a fruitful interplay between the axiomatic and the semantic approach, these two approaches in modal logic remained largely separate. (Cf. the many purely axiomatic papers on modal logic in the *Notre Dame Journal of Formal Logic.*) Two explanations may be adduced for this. First, the algebraic approach to modal logic — itself a generalization of the *truth table* semantics for propositional logic — may have started from the wrong analogy. In many respects, even propositional modal logic behaves like *predicate* logic, whence a generalization of the Tarski semantics (involving interpretation in *structures*) is more appropriate. But, only as late as 1959, Kripke gave such a semantics (cf.[43]). (There were some precursors, such as S. Kanger, cf.[40].) Secondly, and this is a general point, an algebraic semantics does not really provide a "mental picture": one could almost say it is just "syntax in disguise". This does not make it technically useless (on the contrary), but it does make it less appealing intuitively/philosophically. For quite a while after 1960, Kripke semantics held the field; but, as was remarked in the introduction, there is a notable revival of algebraic methods in current research. The "incompleteness theorems" of Fine and Thomason (cf. [22] and [75]) have shown that Kripke semantics as originally intended is essentially weaker than algebraic semantics, so this revival has a clear *rationale.*

Contemporary modal logic, often considered to be a prototype of what may be called *intensional logic*, continues to be a matter of philosophical interest. It will suffice to list some distinguished authors: Quine [65], Hintikka [35], D. Lewis [50], Kripke [44] and Plantinga [63]. Moreover, formally similar philosophical subjects like *deontic logic* (von Wright, cf. [34]) or *tense logic* (cf. Prior [64]) should be included as well. Apart from these philosophical connections, modal logic has been applied in linguistics, as may be seen from the work of R. Montague (cf.[59]), M. J. Cresswell (cf. [18]) and D. Gabbay (cf. [27]). Here we wish to draw attention to some mathematical applications, however. In an early paper (cf. [28]), K. Gödel showed how Heyting's intuitionistic calculus could be embedded in the modal logic *S4*. Gödel's translation enabled Kripke to formulate his semantics for intuitionistic logic as well (cf. [45]). An extensive survey of this area will be found in Smorynski [73]. In fact, intuitionistic research on so-called *intermediate logics* and research on modal logics have much in common (cf. also Blok [14]).

A second important development is the treatment of arithmetical provability predicates as modal necessity operators. This approach, of which R. Montague's paper [60] is an early example, found its present culmination in a paper [74] by R. Solovay showing that the modal logic axiomatized by the formula $L(Lp \to p) \to Lp$ (itself a modalized version of "Löb's Theorem", cf. [53]) contains — so to speak — all the properties of the provability predicate of Peano Arithmetic. This subject has also been studied by the "Siena school", witness the 1975 issue (vol. 34) of *Studia Logica*.

It remains to state the connection between modal logic and second-order logic mentioned in the introduction.

In [77], S. K. Thomason gave an effective translation from the monadic second-order language with unary predicate variables and one, binary predicate constant (i.e., the above language L_2) to modal formulas such that semantic consequence in the former reduces to semantic consequence in the latter. More precisely, there exists a recursive function t from L_2-sentences φ to modal formulas $t(\varphi)$, as well as a special fixed modal formula δ such that, for all sets Σ of L_2-sentences,

$$\Sigma \models \varphi \quad \text{if and only if} \quad \{\delta\} \cup \{t(\sigma) \mid \sigma \in \Sigma\} \models_f t(\varphi).$$

(Here "$\models$" denotes the ordinary relation of second-order semantic consequence on the class of all standard models, and "$\models_f$" denotes modal consequence on the class of all frames, as defined in chapter 2 below.)

Note that L_2 is an extremely powerful language, in which, e.g., second-order Zermelo-Fraenkel set theory may be axiomatized by a single universal sentence ZF^2. This theory contains the complete theory of arithmetic; by the standard set-theoretic reduction of the arithmetical language and the fact that models of ZF^2 contain only the *standard* natural numbers. But, then, it follows from Tarski's theorem on the arithmetical undefinability of arithmetical truth that the notion $\models$ is not arithmetically definable (as a relation between Gödel numbers of sentences, that is); let alone recursively axiomatizable. (For, an arithmetical sentence is true if and only if the first-order sentence in $\in$ which is its set-theoretic reduction follows semantically from ZF^2.) Accordingly, the relation $\models_f$ cannot be recursively axiomatizable either: *standard modal consequence is not axiomatizable*.

The strength of L_2 may also be gauged from an unpublished result of H. C. Doets which shows that the full language of second-order logic (and even of the theory of finite types) is reducible to L_2 in the following sense (cf. [19]). There exists an effective translation **T** from arbitrary second-order sentences φ to second-order sentences $\mathbf{T}(\varphi)$ of the form $\forall R\, \exists P\, \varphi'(R, P)$ such that

$$\models \varphi \quad \text{if and only if} \quad \models \forall R\, \exists P\, \varphi'(R, P).$$

(Here "$\models$" denotes validity on the class of all standard models, R is a binary predicate variable, P is a unary one, and φ' is a first-order sentence in R and P.) The idea of the proof is first to find a first-order sentence $\psi(R, P)$ such that $\forall P \psi(\in, P)$ defines those layers V_α of the set-theoretic cumulative hierarchy for which α is a limit ordinal greater than ω. This may be combined with the well-known set-theoretic (first-order) definition τ of satisfaction for second-order (or also higher-order) sentences. I.e., for all such sentences χ (with set-theoretic "Gödel-codes" $\ulcorner\chi\urcorner$) and all models $\mathbf{A} \in V_\alpha$ of the appropriate type,

$$V_\alpha \models \tau\, (\in, \mathbf{A}, \ulcorner\chi\urcorner) \quad \text{if and only if} \quad \mathbf{A} \models \chi.$$

Thus, the above reduction assumes the following form:

$$\begin{aligned} &\models \chi \quad \text{if and only if} \\ &\models \forall R(\forall P\, \psi(R, P) \rightarrow \forall x\, \tau(R, x, \ulcorner\chi\urcorner)) \quad \text{if and only if} \\ &\models \forall R\, \exists P(\psi(R, P) \rightarrow \forall x\, \tau(R, x, \ulcorner\chi\urcorner)). \end{aligned}$$

It follows from Thomason's and Doets' results taken together that modal formulas serve, in a sense, as a *reduction class* for the theory of finite types; since, for any sentence φ of the latter language,

$$\begin{aligned} \models \varphi \quad & \text{if and only if} \quad \models \mathbf{T}(\varphi) \ (= \forall R\, \exists P\, \varphi'(R, P)) \\ & \text{if and only if} \quad \models \exists P\, \varphi'(R, P) \\ & \text{if and only if} \ \{\delta\} \models_f t(\exists P\, \varphi'(R, P)). \end{aligned}$$

The fixed modal formula δ is indispensable in this result. For, as will be shown in chapter 2 below, universal validity for modal formulas is a recursive notion, whereas, of course, universal validity for higher-order sentences is not. (To see this, note e.g. that in the above example, $ZF^2 \models \varphi$ if and only if the second-order sentence $ZF^2 \rightarrow \varphi$ is universally valid.)

No further general background than the above will be given here. The reader is referred to Hughes & Cresswell [36] for a good introduction, to Segerberg [67] for a solid work on completeness theory, and to Goldblatt [30] for an important monograph on "modal algebra".

CHAPTER II

POSSIBLE WORLDS SEMANTICS

Semantic structures for the modal propositional language are *frames* $\mathbf{F}$ consisting of a non-empty domain W together with a binary relation R on W. Thus, $\mathbf{F} = \langle W, R \rangle$ (and also $\mathbf{F}_1 = \langle W_1, R_1 \rangle$, etc.). In addition to these, auxiliary structures $\mathbf{M}$, called *models*, will be used, consisting of a frame $\mathbf{F}$ together with a *valuation* V *on* $\mathbf{F}$ assigning subsets of W to proposition letters. Thus, $\mathbf{M} = \langle \mathbf{F}, V \rangle$, sometimes also thought of as $\langle W, R, V \rangle$, (and $\mathbf{M}_1 = \langle \mathbf{F}_1, V_1 \rangle$, etc.). Finally, so-called *general frames* will be needed, but these will be defined at a more convenient point (in chapter 4).

The basic truth definition (due to S. Kripke) is as follows.

2.1 *Definition.* Let $\mathbf{M}$ $(= \langle W, R, V \rangle)$ be a model, $w \in W$, and let φ be a modal formula. $\mathbf{M} \models \varphi[w]$ ("φ *holds/is true at* w in $\mathbf{M}$") is defined by recursion through the clauses

(i) $\mathbf{M} \models p[w]$ iff $w \in V(p)$, for any proposition letter p
(ii) $\mathbf{M} \models \neg\varphi[w]$ iff *not* $\mathbf{M} \models \varphi[w]$
(iii) $\mathbf{M} \models (\varphi \wedge \psi)[w]$ iff $\mathbf{M} \models \varphi[w]$ *and* $\mathbf{M} \models \psi[w]$
(iv) $\mathbf{M} \models L\varphi[w]$ iff *for all* $v \in W$ *such that* Rwv: $\mathbf{M} \models \varphi[v]$.

Another choice of primitives would yield clauses

(v) $\mathbf{M} \models M\varphi[w]$ iff *for some* $v \in W$ *such that* Rwv: $\mathbf{M} \models \varphi[v]$
(vi) $\mathbf{M} \models (\varphi \rightarrow \psi)[w]$ iff *if* $\mathbf{M} \models \varphi[w]$, *then* $\mathbf{M} \models \psi[w]$
(vii) $\mathbf{M} \models \bot[w]$ for *no* $\mathbf{M}$ and w.

Using definition 2.1 several derived concepts may be defined:

(a) $\mathbf{M} \models \Sigma[w]$, for a set Σ of modal formulas, if $\mathbf{M} \models \sigma[w]$ for each $\sigma \in \Sigma$
(b) $\mathbf{M} \models \varphi$ if, for all $w \in W$, $\mathbf{M} \models \varphi[w]$
(c) $\mathbf{M} \models \Sigma$ if, for each $\sigma \in \Sigma$, $\mathbf{M} \models \sigma$
(d) $Th_{mod}(\mathbf{M}) = \{\varphi \mid \mathbf{M} \models \varphi\}$
(e) $MOD(\varphi) = \{\mathbf{M} \mid \mathbf{M} \models \varphi\}$
(f) $MOD(\Sigma) = \{\mathbf{M} \mid \mathbf{M} \models \Sigma\}$.

Moreover, the following abbreviation will be useful: "$\mathbf{M} \not\models \varphi[w]$" stands for "*not* $\mathbf{M} \models \varphi[w]$"; and "$\mathbf{M} \not\models \varphi$", etc. have a similar meaning. This "stroke negation" extends easily to a whole range of notations, and will often be used without further announcement.

A modal formula will be called *universally valid* ($\models \varphi$) if it holds in all models at every world w.

The next definition stipulates a direct "contact" between modal formulas and frames.

2.2 *Definition.* Let $\mathbf{F}$ ($= \langle W, R\rangle$) be a frame, $w \in W$, and let φ be a modal formula. $\mathbf{F} \models \varphi[w]$ ("*φ holds/is true at w in* $\mathbf{F}$") if, for all valuations V on $\mathbf{F}$, $\langle \mathbf{F}, V\rangle \models \varphi[w]$.

Again, $\mathbf{F} \models \Sigma[w]$, $\mathbf{F} \models \varphi$, $\mathbf{F} \models \Sigma$, $Th_{mod}(\mathbf{F})$, $FR(\varphi)$ and $FR(\Sigma)$ may be defined in the obvious way.

Another way to phrase these definitions would be by defining $V(\varphi)$ for all formulas φ, starting with the values $V(p)$, through a recursion similar to the one given above. This notation will be used ocasionally, and $V(\varphi)$ is thus defined as $\{w \in W \mid \langle \mathbf{F}, V\rangle \models \varphi[w]\}$. It will always be clear from the context which $\mathbf{F}$ is meant.

Several simple results will illustrate these definitions.

2.3 *Lemma (Isomorphism Lemma).* If f is an isomorphism from the frame $\mathbf{F}_1$ onto the frame $\mathbf{F}_2$, then, for all $w \in W_1$ and all modal formulas φ, $\mathbf{F}_1 \models \varphi[w]$ if and only if $\mathbf{F}_2 \models \varphi[f(w)]$.

2.4 *Lemma (Finiteness Lemma).* If φ contains at most the proposition letters $p_1, \ldots, p_n$, and V_1, V_2 are two valuations on $\mathbf{F}$ such that $V_1(p_i) = V_2(p_i)$ for each i $(1 \leqslant i \leqslant n)$, then, for any $w \in W$,
$\langle \mathbf{F}, V_1\rangle \models \varphi[w]$ if and only if $\langle \mathbf{F}, V_2\rangle \models \varphi[w]$.

2.5 *Lemma (Substitution Lemma).* If $[\psi_1/p_1, ..., \psi_n/p_n]\varphi$ denotes the result of simultaneously substituting ψ_i for p_i $(1 \leqslant i \leqslant n)$ in φ, then, for any model $\langle \mathbf{F}, V\rangle$ and $w \in W$, $\langle \mathbf{F}, V\rangle \models [\psi_1/p_1, ..., \psi_n/p_n]\ \varphi[w]$ if and only if $\langle \mathbf{F},\ V(p_1/V(\psi_1), ..., p_n/V(\psi_n))\rangle \models \varphi[w]$; where $V(p_1/V(\psi_1), ..., p_n/V(\psi_n))$ is like V except for the (possible) difference that the value $V(\psi_i)$ is assigned to p_i $(1 \leqslant i \leqslant n)$.

The next lemma is less obvious. To state it, two important concepts have to be defined first.

2.6 *Definition.* The *degree* $d(\varphi)$ of a modal formula φ is the maximum length of a nested sequence of modal operators in φ, or, inductively,

(i) $d(p) = 0$ *for proposition letters* p
(ii) $d(\neg\varphi) = d(\varphi)$
(iii) $d(\varphi \wedge \psi) = \text{maximum } (d(\varphi), d(\psi))$
(iv) $d(L\varphi) = d(\varphi) + 1$.

2.7 *Definition.* Let $\mathbf{F}$ $(= \langle W, R\rangle)$ be a frame with $w \in W$. $S_n(\mathbf{F}, w)$ ("the *n-hull around* w") is defined by recursion on n:

(i) $S_0(\mathbf{F}, w) = \{w\}$
(ii) $S_{n+1}(\mathbf{F}, w) = S_n(\mathbf{F}, w) \cup \{v \in W \mid \text{for some } u \in S_n(\mathbf{F}, w), Ruv\}$.

2.8 *Lemma.* For any frame $\mathbf{F}$, any $w \in W$ and any two valuations V_1, V_2 on $\mathbf{F}$ such that $V_1(p) \cap S_n(\mathbf{F}, w) = V_2(p) \cap S_n(\mathbf{F}, w)$ for all proposition letters p, it holds that, for any modal formula φ with $d(\varphi) \leqslant n$, $\langle \mathbf{F}, V_1\rangle \models \varphi[w]$ if and only if $\langle \mathbf{F}, V_2\rangle \models \varphi[w]$.

This lemma may be used to show that, for modal formulas, the two primitives $\rightarrow$ and M suffice as far as frames are concerned:

2.9 *Corollary.* For any modal formula φ there exists a modal formula φ_1 whose only logical constants are $\rightarrow$ and M, such that, for all frames $\mathbf{F}$ $(= \langle W, R\rangle)$ and $w \in W$, $\mathbf{F} \models \varphi[w]$ if and only if $\mathbf{F} \models \varphi_1[w]$.

Proof. If $d(\varphi) = 0$, then φ is just a propositional formula and, either φ is a tautology (whence φ_1 may be taken to be $(p \rightarrow p)$), or φ_1 is not a tautology (whence, for no $\mathbf{F}$ and w, $\mathbf{F} \models \varphi[w]$, and φ_1 may be taken to be p). If $d(\varphi) > 0$, then φ_1 is obtained as follows. Rewrite φ to an equivalent formula φ_2 containing only M, $\bot$ and $\rightarrow$. The required φ is, then, $(M^{d(\varphi_2)}q \rightarrow q) \rightarrow (...((M^1q \rightarrow q) \rightarrow (([q/\bot]\varphi_2 \rightarrow q) \rightarrow q))...)$, where $M^i\varphi =_{def} M$ ... (i times) ... $M\varphi$ and q is any proposition letter not occurring in φ_2.

An example of this procedure is provided by $\varphi = M\neg p \rightarrow \neg Mp$, $\varphi_2 = M(p \rightarrow \bot) \rightarrow (Mp \rightarrow \bot)$ and $\varphi_1 = (Mq \rightarrow q) \rightarrow (((M(p \rightarrow q) \rightarrow (Mp \rightarrow q)) \rightarrow q) \rightarrow q)$.

The reason for this definition will become clear from the following argument. (Recall already the tautology $((\varphi \rightarrow \psi) \rightarrow \psi) \leftrightarrow (\varphi \vee \psi)$.) First, note that φ and φ_2 may be taken to be strongly equivalent in the sense that, for all models $\mathbf{M}$ and all w, $\mathbf{M} \models \varphi[w]$ iff $\mathbf{M} \models \varphi_2[w]$. Now, if $\mathbf{F} \models \varphi_1[w]$, then, for any valuation V on $\mathbf{F}$ assigning the empty set to q, $\langle \mathbf{F}, V\rangle \models \varphi_1[w]$. In other words: $\mathbf{F} \models [\bot/q]\varphi_1[w]$. But then, since $M^i\bot \rightarrow \bot$ is universally valid, $\mathbf{F} \models [\bot/q]\ (([q/\bot]\varphi_2 \rightarrow q) \rightarrow q)[w]$, i.e. $\mathbf{F} \models (\varphi_2 \rightarrow \bot) \rightarrow \bot[w]$, so $\mathbf{F} \models \varphi_2[w]$ and hence $\mathbf{F} \models \varphi[w]$. To establish the converse, assume that $\mathbf{F} \models \varphi[w]$. Moreover, let V be a valuation on $\mathbf{F}$ such that $\langle \mathbf{F}, V\rangle \models M^iq \rightarrow q[w]$ for each i $(1 \leqslant i \leqslant d(\varphi_2))$. It is to be shown that $\langle \mathbf{F}, V\rangle \models ([q/\bot]\varphi_2 \rightarrow q) \rightarrow q[w]$. Now, either $\langle \mathbf{F}, V\rangle \models q[w]$ (in which case there is nothing to prove), or $\langle \mathbf{F}, V\rangle \not\models q[w]$, in which case it is to be proven that $\langle \mathbf{F}, V\rangle \models [q/\bot]\varphi_2[w]$. Note that, since $\langle \mathbf{F}, V\rangle \not\models q[w]$, it is not the case that $\langle \mathbf{F}, V\rangle \models M^iq[w]$ for any i $(1 \leqslant i \leqslant d(\varphi_2))$. In other words, $V(q) \cap S_{d(\varphi_2)}(\mathbf{F}, w) = \varnothing$. Lemma 2.8 then implies that $\langle \mathbf{F}, V\rangle \models [q/\bot]\varphi_2[w]$ iff $\langle \mathbf{F}, V(q/\varnothing)\rangle \models [q/\bot]\varphi_2[w]$. But this last statement holds if and only if $\langle \mathbf{F}, V\rangle \models \varphi_2[w]$. And this again follows from the fact that $\langle \mathbf{F}, V\rangle \models \varphi[w]$. **QED.**

The basic notions concerning frames are introduced next.

2.10 *Definition.* A frame $\mathbf{F}_1$ $(= \langle W_1, R_1\rangle)$ is a *subframe* of a frame $\mathbf{F}_2$ $(= \langle W_2, R_2\rangle)$ (notation: $\mathbf{F}_1 \subseteq \mathbf{F}_2$) if (i) $W_1 \subseteq W_2$ and (ii) $R_1 = R_2 \cap (W_1 \times W_1)$. $\mathbf{F}_1$ is a *generated subframe* of $\mathbf{F}_2$ (notation: $\mathbf{F}_1 \underset{\rightarrow}{\subseteq} \mathbf{F}_2$) if $\mathbf{F}_1$ is a subframe of $\mathbf{F}_2$ such that, for all $w \in W_1$, $v \in W_2$, R_2wv only if $v \in W_1$. A model $\mathbf{M}_1$ $(= \langle \mathbf{F}_1, V_1\rangle)$ is a *submodel* of a model $\mathbf{M}_2$ $(= \langle \mathbf{F}_2, V_2\rangle)$ (notation: $\mathbf{M}_1 \subseteq \mathbf{M}_2$) if $\mathbf{F}_1$ is a subframe of $\mathbf{F}_2$ and, for each proposition letter p, $V_1(p) = V_2(p) \cap W_1$. If, in addition, $\mathbf{F}_1$ is a generated subframe of $\mathbf{F}_2$, then $\mathbf{M}_1$ is a *generated submodel* of $\mathbf{M}_2$ (notation: $\mathbf{M}_1 \underset{\rightarrow}{\subseteq} \mathbf{M}_2$).

The notion "generated subframe" is closely related to the better-known notion "end-extension" (cf. Chang & Keisler [17]).

2.11 *Lemma* (*Generation Theorem*, Segerberg [69], Feferman [21]). If $\mathbf{M}_1 \underset{\rightarrow}{\subseteq} \mathbf{M}_2$, then, for all $w \in W_1$ and all modal formulas φ, $\mathbf{M}_2 \models \varphi[w]$ iff $\mathbf{M}_1 \models \varphi[w]$.

A kind of converse to 2.11 holds as well. Let $\mathbf{F}_1$ be a subframe of $\mathbf{F}_2$. Given a valuation V_2 on $\mathbf{F}_2$, V_1 is the valuation on $\mathbf{F}_1$ defined by $V_1(p) = V_2(p) \cap W_1$ for all proposition letters p. Then, if, for all valuations V_2 on $\mathbf{F}_2$, all $w \in W_1$ and all modal formulas φ, $\langle \mathbf{F}_2, V_2 \rangle \models \varphi[w]$ iff $\langle \mathbf{F}_1, V_1 \rangle \models \varphi[w]$, $\mathbf{F}_1$ is a generated subframe of $\mathbf{F}_2$.

2.12 *Corollary.* If $\mathbf{F}_1 \underset{\rightarrow}{\subseteq} \mathbf{F}_2$, then, for all $w \in W_1$ and all modal formulas φ,

$\mathbf{F}_2 \models \varphi[w]$ iff $\mathbf{F}_1 \models \varphi[w]$

$\mathbf{F}_2 \models \varphi$ only if $\mathbf{F}_1 \models \varphi$ (i.e. modal formulas are *preserved under generated subframes*).

Corollary 2.12 implies that the L_0-sentence $\exists x \exists y Rxy$ is not defined by any modal formula — it is not preserved under generated subframes.

An important kind of generated subframe is the following.

2.13 *Definition.* Let $\mathbf{F}$ be a frame and $w \in W$. The *subframe* of $\mathbf{F}$ *generated by* w ($TC(\mathbf{F}, w)$) is the smallest generated subframe $\mathbf{F}_1$ ($= \langle W_1, R_1 \rangle$) of $\mathbf{F}$ such that $w \in W_1$. I.e. $W_1 = \bigcap \{X \subseteq W \mid w \in X$ and, for all $v \in X$ and $u \in W$, Rvu only if $u \in X\}$ $= \{v \in W \mid$ a sequence $v_1, \ldots, v_n$ exists with $w = v_1$, $v = v_n$ and Rv_iv_{i+1} for each i $(1 \leqslant i \leqslant n-1)$. A frame $\mathbf{F}$ of the form $TC(\mathbf{F}, w)$ for some $w \in W$ is called *generated*.

Clearly, $\mathbf{F} \models \varphi[w]$ iff $TC(\mathbf{F}, w) \models \varphi[w]$.

The next notion constructs one frame out of many.

2.14 *Definition.* Let $\{\mathbf{F}_i \mid i \in I\}$ be a family of frames. Set $\mathbf{F}_i' = \langle W_i', R_i' \rangle$, where $W_i' = \{\langle i, w \rangle \mid w \in W_i\}$ and $R_i' = \{\langle \langle i, w \rangle, \langle i, v \rangle \rangle \mid \langle w, v \rangle \in R_i\}$. The *disjoint union* $\Sigma \{\mathbf{F}_i \mid i \in I\}$ of the family $\{\mathbf{F}_i \mid i \in I\}$ is the frame $\langle \bigcup \{W_i' \mid i \in I\}, \bigcup \{R_i' \mid i \in I\} \rangle$.

For each $i \in I$, F_i is isomorphic to the generated subframe F_i' of $\Sigma\{\mathsf{F}_i \mid i \in I\}$, whence the following corollary.

2.15 *Corollary.* For each $i \in I$, $w \in W_i$ and all modal formulas φ, $\mathsf{F}_i \models \varphi[w]$ iff $\Sigma\{\mathsf{F}_i \mid i \in I\} \models \varphi[\langle i, w\rangle]$; whence $\Sigma\{\mathsf{F}_i \mid i \in I\} \models \varphi$ iff, for each $i \in I$, $\mathsf{F}_i \models \varphi$.

The fact that if, for each $i \in I$, $\mathsf{F}_i \models \varphi$, then $\Sigma\{\mathsf{F}_i \mid i \in I\} \models \varphi$, is expressed by saying that modal formulas are ***preserved under disjoint unions***. Corollary 2.15 implies that $\forall x \forall y Rxy$ — although preserved under generated subframes — is not defined by any modal formula, not being preserved under disjoint unions.

2.16 *Definition.* A function f from a frame F_1 to a frame F_2 is a *p-morphism* if (i) for all $w, v \in W_1$, R_1wv only if $R_2f(w)f(v)$ (i.e. if f is a *homomorphism*) and (ii) for all $w \in W_1$ and $v \in W_2$, $R_2f(w)v$ only if there exists a $u \in W_1$ such that R_1wu and $f(u) = v$. f is a ***p-morphism*** from a model $\langle \mathsf{F}_1, V_1\rangle$ to a model $\langle \mathsf{F}_2, V_2\rangle$ if it is a p-morphism from F_1 to F_2 and also, for each $w \in W_1$ and each proposition letter p, $w \in V_1(p)$ iff $f(w) \in V_2(p)$.

Note that any 1-1 *onto* p-morphism is an isomorphism. The concept of a "p-morphism" was first defined by K. Segerberg in [69]. An earlier similar notion ("strongly isotone function") is in De Jongh & Troelstra [39].

2.17 *Lemma* (*p-morphism theorem*, Segerberg [69]). If f is a p-morphism from M_1 onto M_2, then, for all $w \in W_1$ and all modal formulas φ, $\mathsf{M}_1 \models \varphi[w]$ iff $\mathsf{M}_2 \models \varphi[f(w)]$.

A kind of converse to 2.17 holds as well. Let f be any function from F_1 onto F_2. Given a valuation V_2 on F_2, $f^{-1}(V_2)$ is the valuation on F_1 defined by $f^{-1}(V_2)(p) = f^{-1}[V_2(p)]$. Then, if, for all valuations V_2 on F_2, all $w \in W_1$ and all modal formulas φ, $\langle \mathsf{F}_1, f^{-1}(V_2)\rangle \models \varphi[w]$ iff $\langle \mathsf{F}_2, V_2\rangle \models \varphi[f(w)]$, then f is a p-morphism.

2.18 *Corollary.* If f is a p-morphism from F_1 onto F_2, then, for all $w \in W_1$ and all modal formulas φ,

$\mathbf{F}_1 \models \varphi[w]$ only if $\mathbf{F}_2 \models \varphi[f(w)]$
$\mathbf{F}_1 \models \varphi$ only if $\mathbf{F}_2 \models \varphi$ (i.e. modal formulas are *preserved under p-morphic images).*

It follows that $\forall x \neg Rxx$, although preserved under generated subframes and disjoint unions, is not defined by any modal formula. For, it is not preserved under the p-morphism f from $\langle IN, < \rangle$ onto $\langle \{0\}, \{\langle 0, 0\rangle\}\rangle$ defined by $f(n) = 0$ for all $n \in IN$.

Lemma 2.17 has several interesting consequences.

2.19 *Lemma* (cf. Makinson [57]). For any frame $\mathbf{F}$ and any modal formula φ, if $\mathbf{F} \models \varphi$, then $\langle \{0\}, \{\langle 0, 0\rangle\}\rangle \models \varphi$ or $\langle \{0\}, \varnothing\rangle \models \varphi$.

Proof. If W contains an element w without R-successors, then $\langle \{w\}, \varnothing\rangle \underset{\rightarrow}{\subseteq} \mathbf{F}$. Therefore, by corollary 2.12, $\langle \{w\}, \varnothing\rangle \models \varphi$.

If every element of W has an R-successor, then f defined by $f(w) = 0$ for all $w \in W$ is a p-morphism from $\mathbf{F}$ onto $\langle \{0\}, \{\langle 0, 0\}\rangle$, whence, by corollary 2.18, $\langle \{0\}, \{\langle 0, 0\rangle\}\rangle \models \varphi$. QED.

2.20 *Lemma.* Any frame $\mathbf{F}$ is a p-morphic image of some disjoint union of generated frames.

Proof. The following stipulation defines a p-morphism from $\Sigma\{TC(\mathbf{F}, w) \mid w \in W\}$ onto $\mathbf{F}$: $f(\langle w, v\rangle) = v$. QED.

2.21 *Lemma* (*Tree Lemma*, cf. Sahlqvist [66]). Any generated frame $\mathbf{F}$ is a p-morphic image of some irreflexive intransitive tree in which no R-loops occur.

Proof. Let $\mathbf{F} = TC(\mathbf{F}, w)$, for $w \in W$. Define a tree whose nodes are the finite sequences $\langle w_1, \ldots, w_n\rangle$ $(w_1, \ldots, w_n \in W)$ satisfying $Rw_1w_2, \ldots, Rw_{n-1}w_n$, and whose successor relation joins nodes $\langle w_1, \ldots, w_n\rangle$ and $\langle w_1, \ldots, w_n, w_{n+1}\rangle$ with Rw_nw_{n+1}. The map f defined by $f(\langle w_1, \ldots, w_n\rangle) = w_n$ is a p-morphism from this tree onto $\mathbf{F}$. QED.

2.22 *Corollary.* Any modal formula φ which is not universally valid is falsified on some finite irreflexive intransitive tree.

Proof. If φ is not universally valid, then, for some $\mathbf{F}$ and w, $\mathbf{F} \not\models \varphi[w]$.

It follows, by 2.12, that $TC(\mathbf{F}, w) \not\models \varphi[w]$. Therefore, by 2.21 and 2.18, φ is falsified on some irreflexive intransitive tree. The next lemma implies that a finite subtree of this tree suffices for falsifying φ. **QED.**

2.23 *Lemma.* Let $\mathbf{F}$ be an irreflexive intransitive tree in which no R-loops occur, V a valuation on $\mathbf{F}$ and $w \in W$. Let $\varphi_1, \ldots, \varphi_n, \psi_1, \ldots, \psi_m$ be modal formulas such that $\mathbf{M}$ $(= \langle \mathbf{F}, V \rangle) \models \varphi_i[w]$ for each i $(1 \leqslant i \leqslant n)$, but, for no j $(1 \leqslant j \leqslant m)$, $\mathbf{M} \models \psi_j[w]$. There exists a finite submodel $\mathbf{M}'$ of $TC(\mathbf{M}, w)$ with w in its domain such that $\mathbf{M}' \models \varphi_i[w]$ for each i $(1 \leqslant i \leqslant n)$, but, for no j $(1 \leqslant j \leqslant m)$, $\mathbf{M}' \models \psi_j[w]$.

Proof. This lemma is proven by induction on the number N of occurrences of Boolean and modal operators in $\varphi_1, \ldots, \varphi_n, \psi_1, \ldots, \psi_m$. If $N = 0$, then it suffices to take the submodel $\mathbf{M}'$ with domain $\{w\}$. If the lemma is assumed to hold for numbers up to N, then the case for N is treated as follows. If Boolean reductions are possible, the procedure is trivial. If not, then each φ_n and ψ_j is either a proposition letter, or of the form $L\chi$. Let $L\chi_1, \ldots, L\chi_r$ be the formulas of the second kind occurring among $\psi_1, \ldots, \psi_m$. Choose $w_1, \ldots, w_r \in W$ such that Rww_i, but $\mathbf{M} \not\models \chi_i[w_i]$, for each i $(1 \leqslant i \leqslant r)$. By the induction hypothesis, finite submodels $\mathbf{M}'_1, \ldots, \mathbf{M}'_r$ of $TC(\mathbf{M}, w_1), \ldots, TC(\mathbf{M}, w_r)$ exist, containing $w_1, \ldots, w_r$ respectively, such that, for all formulas $L\varphi$ occurring among $\varphi_1, \ldots, \varphi_n$, $\mathbf{M}'_i \models \varphi[w_i]$, but $\mathbf{M}'_i \not\models \chi_i[w_i]$, for each i $(1 \leqslant i \leqslant r)$. The submodel of $\mathbf{M}$ obtained from the union of $\mathbf{M}'_1, \ldots, \mathbf{M}'_r$ by adding w is the required $\mathbf{M}'$. (To see this, note that w belongs to no n-hull around v for any $v \in W'_1 \cup \ldots \cup W'_r$; because otherwise $\mathbf{F}$ would contain an R-loop. Hence the addition of w does not change truth or falsity of modal formulas in $\mathbf{M}'_1, \ldots, \mathbf{M}'_r$ — by lemma 2.8.) QED.

The next concept concerning frames is motivated by the algebraic theory of chapter 4.

Recall the notation "$l(X)$" introduced earlier on: for a frame $\mathbf{F} = \langle W, R \rangle$, $X \subseteq W$, $l(X) = \{w \in W \mid$ for all $v \in W$, Rwv only if $v \in X\}$.

2.24 *Definition.* The *ultrafilter extension* $ue(\mathbf{F})$ of a frame $\mathbf{F}$ $(= \langle W, R \rangle)$ is the frame $\langle W_{\mathbf{F}}, R_{\mathbf{F}} \rangle$ whose domain W consists of all ultrafilters on W and whose relation $R_{\mathbf{F}}$ consists of those pairs of ultrafilters U_1, U_2 such that, for all $X \subseteq W$, if $l(X) \in U_1$, then $X \in U_2$.

2.25 *Lemma.* If V is a valuation on $\mathbf{F}$ and $ue(V)$ the valuation on $ue(\mathbf{F})$ defined by $ue(V)(p) = \{U \mid V(p) \in U\}$, then, for all ultrafilters U on W and all modal formulas φ, $\langle ue(\mathbf{F}), ue(V)\rangle \models \varphi[U]$ iff $V(\varphi) \in U$.

Proof. The proof is by induction on the complexity of φ. The case $\varphi = p$ is trivial, using the definition of $ue(V)$. The cases $\varphi = \neg\psi$ and $\varphi = \psi \wedge \chi$ are standard, using the defining properties of ultrafilters. Consider then the case $\varphi = L\psi$. If $V(L\psi)$ $(= l(V(\psi))) \in U$, then, by the definition of $R_{\mathbf{F}}$, $V(\psi) \in U_1$ for all U_1 such that $R_{\mathbf{F}}UU_1$. Therefore, by the induction hypothesis, $\langle ue(\mathbf{F}), ue(V)\rangle \models \psi[U_1]$ in such U_1's, whence $\langle ue(\mathbf{F}), ue(V)\rangle \models L\psi[U]$. It remains to establish the converse direction, which is less trivial.

Suppose that $V(L\psi) \in U$. An ultrafilter U_1 is to be found such that $R_{\mathbf{F}}UU_1$ and $\langle ue(\mathbf{F}), ue(V)\rangle \not\models \psi[U_1]$; i.e., by the induction hypothesis, $V(\psi) \notin U_1$. To get such a U_1, it suffices to prove that the set $\{X \subseteq W \mid l(X) \in U\} \cup \{W - V(\psi)\}$ has the finite intersection property. (For in that case, the fundamental theorem on ultrafilters may be applied to extend it to an ultrafilter U_1 not containing $V(\psi)$, for which it is evident that, if $l(X) \in U$, then $X \in U_1$.) Suppose, therefore, that for some $X_1, ..., X_n$ as described, $X_1 \cap ... \cap X_n \cap (W - V(\psi)) = \emptyset$, i.e., $X_1 \cap ... \cap X_n \subseteq V(\psi)$. Then $l(X_1 \cap ... \cap X_n) = l(X_1) \cap ... \cap l(X_n) \subseteq l(V(\psi)) = V(L\psi)$. But, this intersection is in U, whence $V(L\psi)$ would be in U, contradicting the assumption that $V(L\psi) \notin U$. QED.

2.26 *Corollary.* For any frame $\mathbf{F}$ and any modal formula φ, $ue(\mathbf{F}) \models \varphi$ only if $\mathbf{F} \models \varphi$.

It follows that the L_0-sentence $\forall x \exists y(Rxy \wedge Ryy)$ — although preserved under generated subframes, disjoint unions and p-morphisms — is not defined by any modal formula. For, its negation is not preserved under ultrafilter extensions. This is shown by the example of $\mathbf{F} = \langle IN, < \rangle$, in which its negation holds, whereas this formula itself holds in $ue(\langle IN, <\rangle)$. To see this, note that, for any ultrafilter U_1 and any free ultrafilter U_2 on IN, $R_{\mathbf{F}}U_1U_2$ holds. (For, if $l(X) \in U_1$, then X must be cofinite; whence $X \in U_2$, U_2 being free.) In particular, this implies that $R_{\mathbf{F}}UU$ for all free ultrafilters U. Hence, clearly $\forall x \exists y(Rxy \wedge Ryy)$ holds on $ue(\langle IN, <\rangle)$. (This proof is due to R. I. Goldblatt & S. K. Thomason [31].)

Not all modal formulas are preserved under ultrafilter extensions. A counterexample is provided by "Löb's Formula" $L(Lp \rightarrow p) \rightarrow Lp$ (LF). As will be shown in chapter 3, LF holds on exactly the frames $\mathbf{F}$ ($= \langle W, R\rangle$) for which R is transitive and the converse relation of R is well-founded. E.g. $\langle IN, >\rangle \models LF$. But, by an argument similar to the above, $ue(\langle IN, >\rangle)$ turns out to contain reflexive elements; whence the converse of its relation is not well-founded, and LF fails to hold. In chapters 6, 13 and 16, it will be seen, however, that the modal formulas which are preserved under ultrafilter extensions (the so-called *canonical* modal formulas) form a very important class.

2.27 *Lemma*. Any frame $\mathbf{F}$ is isomorphically embedded as a subframe of $ue(\mathbf{F})$.

Proof. The embedding f is given by $f(w) = \{X \subseteq W \mid w \in X\}$. QED.

$\mathbf{F}$ need not be isomorphic to a *generated* subframe of $ue(\mathbf{F})$, however, as was shown by the example of $\forall x \exists y(Rxy \wedge Ryy)$. It holds in $ue(\langle IN, <\rangle)$ and it is preserved under generated subframes, but it does not hold in $\langle IN, <\rangle$.

The main concepts introduced up to now (*generated subframe*, *disjoint union*, *p-morphic image* and *ultrafilter extension*) have a strong algebraic motivation. As will be seen in chapter 4, they correspond, in a sense, to the following algebraic concepts: *homomorphic image*, *direct product*, *subalgebra* and *Stone representation*. Moreover, the following theorem (to be proven in chapter 16) shows that they characterize the modal formulas that define first-order relational properties.

2.28 *Theorem* (R. I. Goldblatt & S.K. Thomason [31]). A class of frames which is closed under elementary equivalence is of the form $\{\mathbf{F} \mid \mathbf{F} \models \Sigma\}$ for some set Σ of modal formulas if and only if it is closed under generated subframes, disjoint unions and p-morphic images, while its complement is closed under ultrafilter extensions.

Unfortunately, theorem 2.28 does not hold generally, without the restriction to closure under elementary equivalence (i.e., to $\Sigma\Delta$-elementary classes; cf. chapter 8). This will follow from a counterexample in chapter 3.

Another important concept is that of a *filtration* (cf. Segerberg [69]).

2.29 *Definition.* Let $\mathbf{M}_1$ (= $\langle W_1, R_1, V_1\rangle$) and $\mathbf{M}_2$ (= $\langle W_2, R_2, V_2\rangle$) be models, and let Σ be a set of modal formulas closed under the formation of subformulas. A function g from $\mathbf{M}_1$ onto $\mathbf{M}_2$ is a *filtration with respect to* Σ if the following three conditions are satisfied.

(i) for all $w, v \in W_1$, R_1wv only if $R_2g(w)g(v)$ (i.e., g is a homomorphism)

(ii) for all $w \in W_1$ and all proposition letters p in Σ, $w \in V_1(p)$ iff $g(w) \in V(p)$

(iii) for all $w \in W_1$ and all modal formulas φ such that $L\varphi \in \Sigma$, $\mathbf{M}_1 \models L\varphi[w]$ only if $\mathbf{M}_2 \models L\varphi[g(w)]$.

2.30 *Lemma (Filtration Lemma).* If g is a filtration with respect to Σ from $\mathbf{M}_1$ onto $\mathbf{M}_2$, then, for all $w \in W_1$ and all modal formulas $\varphi \in \Sigma$, $\mathbf{M}_1 \models \varphi[w]$ iff $\mathbf{M}_2 \models \varphi[g(w)]$.

The standard example is the *modal collapse* of a model $\mathbf{M}$ (= $\langle W, R, V\rangle$) with respect to a set of formulas Σ closed under the formation of subformulas. It is defined as the model $\langle W_\Sigma, R_\Sigma, V_\Sigma\rangle$, where, for $g(w) = \{\varphi \in \Sigma \mid \mathbf{M} \models \varphi[w]\}$, $W_\Sigma = g[W]$, $R_\Sigma g(w)g(v)$ if, for all modal formulas φ such that $L\varphi \in \Sigma$, $L\varphi \in g(w)$ only if $\varphi \in g(v)$, $V_\Sigma(p) = \{g(w) \mid p \in g(w)\}$ for all proposition letters p occurring in Σ (all others are disregarded). It is obvious that a function g as defined here is a filtration from $\mathbf{M}$ onto its modal collapse.

The modal collapse of $\mathbf{M}$ with respect to Σ is sometimes called the *greatest filtration* of $\mathbf{M}$ with respect to Σ. The reason for this is that, if g' is any filtration with respect to Σ from $\mathbf{M}$ onto a model $\mathbf{M}'$ whose domain is W_Σ, then $R'g(w)g(v)$ implies $R_\Sigma g(w)g(v)$ for all $w,v \in W$. There exists also a *smallest filtration* of $\mathbf{M}$ with respect to Σ, with W_Σ and V_Σ defined as above, but whose relation R_s is defined by $R_sg(w)g(v)$ if, for some $w_1, v_1 \in W$ such that $g(w_1) = g(w)$ and $g(v_1) = g(v)$, Rw_1v_1. It is easily checked that, if g' is any filtration with respect to Σ from $\mathbf{M}$ onto a model $\mathbf{M}'$ whose domain is W_Σ, then $R_sg(w)g(v)$ implies $R'g(w)g(v)$ for all $w, v \in W$.

All filtrations are homomorphisms, but they need not be p-morphisms. (There are obvious counterexamples to this.) In several important cases, a filtration *is* a p-morphism, however. E.g., when Σ is the set of *all* formulas and the modal collapse with respect to Σ is *finite*, then the collapsing function is a p-morphism. (To see this, note that any element of the modal collapse is, in this case, defined by a single modal formula.) This subject will be not pursued here.

The filtration technique has many applications in modal completeness theory (cf. Segerberg [67]). Here, only two applications will be mentioned, both consisting in new proofs of results already obtained by different means.

Lemma 2.22 implies that any modal formula which is not universally valid is falsified on some finite model. This may be shown very quickly as follows. If φ is not universally valid, then, for some frame $\mathbf{F}$, some $w \in W$ and some valuation V on $\mathbf{F}$, $\langle \mathbf{F}, V\rangle \models \neg\varphi[w]$. The modal collapse of $\langle \mathbf{F}, V\rangle$ with respect to the finite set Σ consisting of $\neg\varphi$ together with all its subformulas, yields a *finite* model $\mathbf{M}$ in which φ is falsified. This property is called the *finite model property*. It implies that the set of modal formulas which are not universally valid is recursively enumerable. (Searching through the finite models will eventually yield all of them.) Now it is easy to see that the set of universally valid modal formulas is recursively enumerable as well (cf. chapters 3 and 5). Therefore, by Post's Theorem, universal validity of modal formulas is a recursive notion.

2.31 *Corollary*. The set of universally valid modal formulas is recursive.

A second application of filtrations concerns the above-mentioned formula $\forall x \exists y(Rxy \wedge Ryy)$. An involved proof (using ultrafilter extensions) established that no modal formula defines it. By using filtrations, an elementary proof becomes possible. Suppose that some set Δ of modal formulas defines $\forall x \exists y(Rxy \wedge Ryy)$. Since this formula fails to hold in $\langle IN, <\rangle$, there exists some $\varphi \in \Delta$ such that, for some valuation V and $n \in IN$, $\langle IN, <, V\rangle \models \neg\varphi[n]$. Let Σ be the finite set consisting of $\neg\varphi$ together with all its subformulas. Define a filtration g from $\langle IN, <, V\rangle$ to a finite model $\mathbf{M}_1 = \langle W_1, R_1, V_1\rangle$ falsifying φ, by setting:

$g(n) = \{\varphi \in \Sigma \mid \langle IN, <, V\rangle \models \varphi[n]\}$
$W_1 = g[IN]$
$R_1 g(m)g(n)$ if, for each formula φ such that $L\varphi \in \Sigma$, if $L\varphi \in g(m)$, then *both* $L\varphi$ *and* φ are in $g(n)$,
$V_1(p) = \{g(m) \mid p \in g(m)\}$ for any proposition letter $p \in \Sigma$.

(This choice of R_1 yields a kind of modal collapse with respect to Σ often referred to as a *Lemmon filtration.*) That g is a filtration with respect to Σ follows from the fact that $<$ is transitive. Moreover, it is easily seen that R_1 is transitive. R_1 also satisfies $\forall x \exists y Rxy$, because $<$ satisfies this condition and g, being a homomorphism, transmits it. But, any finite frame satisfying both $\forall x \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))$ and $\forall x \exists y Rxy$ also satisfies $\forall x \exists y(Rxy \wedge Ryy)$. It follows that Δ should hold in $\langle W_1, R_1\rangle$, contradicting the fact that $\varphi \in \Delta$ is falsified there. QED.

The Lemmon filtration exhibits an important feature of the filtration technique. One often tries to define the relation R (in the collapse) in such a way that certain desirable properties of the original relation (in our case: transitivity) are preserved.

For a final important use of filtration, see the "Correspondence Theory" chapter in the *Handbook of Philosophical Logic* (mentioned in the Introduction). There, the notion of *modal equivalence* is introduced for frames, meaning that they have the same modal theory. One can prove such equivalences by transferring modal "counterexamples" from one frame to the other, by first "filtrating" and then "embedding". Studying the sieve of modal equivalence tells us a lot about the expressive power of our modal language. For certain special cases, such as *well-ordered* frames, complete classifications have been found in the meantime. (Roughly speaking, modal logic only recognizes the countable order types $\omega \cdot n + k$, $\omega \cdot \omega + k$ $(k, n \in \omega)$.)

It remains to define the notion of semantic consequence to which the above gives rise. There are two of these, both splitting up into a *local* and a *global* variant.

2.32 *Definition.* For a set Σ of modal formulas and a modal formula φ,

$\Sigma \models_f \varphi$ if, for all frames $\mathbf{F}$, $\mathbf{F} \models \Sigma$ only if $\mathbf{F} \models \varphi$ (*global consequence*)

$\Sigma \models_{f,l} \varphi$ if, for all frames $\mathbf{F}$ and $w \in W$, $\mathbf{F} \models \Sigma[w]$ only if $\mathbf{F} \models \varphi[w]$ (*local consequence*)

$\Sigma \models_m \varphi$ if, for all models $\mathbf{M}$, $\mathbf{M} \models \Sigma$ only if $\mathbf{M} \models \varphi$

$\Sigma \models_{m,l} \varphi$ if, for all models $\mathbf{M}$ and $w \in W$, $\mathbf{M} \models \Sigma[w]$ only if $\mathbf{M} \models \varphi[w]$.

The global variants are more natural from our point of view. But, if the basic notion of frame were taken to be $\langle \mathbf{F}, w \rangle$, with a distinguished *actual world* w (as in Kripke's original publications), then the local variants would be preferable. Fortunately, the two notions are easily comparable. Let $L(\Sigma)$ be the set $\{L^i\varphi \mid \varphi \in \Sigma, i \geqslant 0\}$.

2.33 *Lemma.* For any set Σ of modal formulas and any modal formula φ,

$$\Sigma \models_f \varphi \text{ iff } L(\Sigma) \models_{f,l} \varphi$$
$$\Sigma \models_m \varphi \text{ iff } L(\Sigma) \models_{m,l} \varphi.$$

Proof. Use lemma 2.11. QED.

Attention will, therefore, be restricted to the global variants. It is easy to show that $\Sigma \models_m \varphi$ implies $\Sigma \models_f \varphi$. The converse implication fails, however. E.g., $\{p\} \models_f \perp$, but $\{p\} \not\models_m \perp$. It will be seen in chapter 6 that a completeness theorem exists for $\models_m$; whereas no such result holds for $\models_f$, in view of S. K. Thomason's result mentioned in chapter 1: $\models_f$ is "too strong".

Note that the difference between local and global variants also occurs in standard logic. There, one has a choice between defining $\Sigma \models \varphi$ "locally", as "for all structures $\mathbf{D}$ and all assignments b, $\mathbf{D} \models \Sigma[b]$ implies $\mathbf{D} \models \varphi[b]$", or "globally", as "for all structures $\mathbf{D}$, $\mathbf{D} \models \Sigma[b]$ for all assignments b only if $\mathbf{D} \models \varphi[b]$ for all assignments b". This terminology differs from mathematical practice in that the local variant is stronger than the global one.

CHAPTER III

DEFINABILITY

When modal formulas are interpreted in *models*, as in definition 2.1, they are equivalent to a special kind of first-order formulas. This will become clear from the following translation taking modal formulas φ to L_1-formulas $ST(\varphi)$ containing one free variable.

3.1 *Definition.* Let x be a fixed individual variable.

(i) $ST(p) = Px$
(ii) $ST(\neg\psi) = \neg ST(\psi)$
(iii) $ST(\psi \wedge \chi) = ST(\psi) \wedge ST(\chi)$
(iv) $ST(L\psi) = \forall y(Rxy \rightarrow [y/x]ST(\psi))$,

where y is the first variable (in some fixed enumeration of the individual variables) not occurring in $ST(\psi)$.

E.g., $ST(Lp \rightarrow p) = \forall y(Rxy \rightarrow Py) \rightarrow Px$,
$ST(Lp \rightarrow LLp) = \forall y(Rxy \rightarrow Py) \rightarrow \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Pz))$.

Supplementary, but redundant, clauses may be added, like

(v) $ST(M\psi) = \exists y(Rxy \wedge [y/x]ST(\psi))$.

Since models are L_1-structures as well, an obvious equivalence holds:

$\mathbf{M} \models \varphi[w]$ iff $\mathbf{M} \models ST(\varphi)[w]$ (where w is assigned to x)
$\mathbf{M} \models \varphi$ iff $\mathbf{M} \models \forall x ST(\varphi)$

By means of this equivalence, well-known results about L_1 become applicable to modal formulas. E.g., one immediately obtains a *Löwenheim-Skolem theorem* for models, a *compactness theorem* (for $\models_m$, that is) and a *completeness theorem* (in the sense that the set of universally valid formulas is recursively axiomatizable).

When modal formulas are interpreted in *frames*, they become second-order formulas (in L_2), however. If the modal formula φ contains the proposition letters $p_1, \ldots, p_n$, then the following equivalence holds:

$$\begin{array}{lll} \mathbf{F} \models \varphi[w] & \text{iff } \mathbf{F} \models \forall P_1 \ldots \forall P_n ST(\varphi)[w] \\ \mathbf{F} \models \varphi & \text{iff } \mathbf{F} \models \forall x \forall P_1 \ldots \forall P_n ST(\varphi) \\ & \text{iff } \mathbf{F} \models \forall P_1 \ldots \forall P_n \forall x ST(\varphi). \end{array}$$

Now, the Löwenheim-Skolem theorem fails: S. K. Thomason has an example in [78] of an uncountable frame $\mathbf{F}$ such that, for no countable frame $\mathbf{F}'$, $Th_{mod}(\mathbf{F}) = Th_{mod}(\mathbf{F}')$. The compactness theorem fails too: the same author has an example in [76] of a set Σ of modal formulas and a modal formula φ such that $\Sigma \models_f \varphi$, but, for no finite $\Sigma_0 \subseteq \Sigma$, $\Sigma_0 \models_f \varphi$. This last result follows from Thomason's reduction of second-order logic to modal logic (cf. chapter 1) combined with the non-compactness of the former logic. Moreover, although universal validity is a recursively axiomatizable notion (as was noted above), the relation $\models_f$ of modal consequence, even between single modal formulas, is not recursively axiomatizable, as was shown in chapter 1.

Modal formulas as (first-order) L_1-formulas will be discussed now. Later on in this chapter, the emphasis will shift to modal formulas as (second-order) L_2-formulas.

There is a more elegant, independent way of describing the L_1-formulas of the form $ST(\varphi)$ for some modal φ:

3.2 *Definition.* An *m-formula* is a member of the smallest class X of L_1-formulas satisfying

(i) $Px \in X$ for each unary predicate constant P and each individual variable x

(ii) if $\alpha \in X$, then $\neg\alpha \in X$

(iii) if $\alpha \in X$ and $\beta \in X$, then $(\alpha \wedge \beta) \in X$

(iv) if $\alpha \in X$ and x, y are distinct individual variables, then $\forall y(Rxy \rightarrow \alpha) \in X$.

Thus, m-formulas are formulas constructed from *unary* atoms using Boolean operators and *restricted* universal quantification. The translations of modal formulas belong to a yet more restricted class:

3.3 *Definition.* An *M-formula* is a member of the smallest class X of L_1-formulas satisfying

(i) $Px \in X$ for each unary predicate constant P and each individual variable x
(ii) if $\alpha \in X$, then $\neg\alpha \in X$
(iii) if α and β have the same free individual variable and are both in X, then $(\alpha \wedge \beta) \in X$
(iv) if $\alpha \in X$ and y is the free variable of α, then $\forall y(Rxy \rightarrow \alpha) \in X$, provided that x is distinct from y.

m-formulas have at least one free variable, *M*-formulas have exactly one.

3.4 *Lemma.* Any *m*-formula α is equivalent to a Boolean combination of *M*-formulas, each with their free variable among those of α.

Proof. The assertion is proved by induction on the complexity of *m*-formulas. In order to simplify the proof, the above definition is changed as follows. A clause for *disjunction* ($\vee$) is added and the clause for restricted universal quantification is replaced by one for restricted *existential* quantification. As we are only trying to prove an equivalence this change is harmless.

The cases $\alpha = Px$, $\alpha = \neg\beta$, $\alpha = \beta \wedge \gamma$ and $\alpha = \beta \vee \gamma$ are trivial. It remains to consider $\alpha = \exists y(Rxy \wedge \beta)$. By the induction hypothesis, β is equivalent to a Boolean combination of *M*-formulas, each with their free variable among those of β. By the theorem on distributive normal forms, β is then equivalent to a formula of the form $\sum_{i=1}^{n} \prod_{j=1}^{n_i} \beta_{ij}$, where each β_{ij} is an *M*-formula. (As for the notation, the relevant stipulation is that $\sum_{i=1}^{n} \varphi_i =_{def} (\varphi_1 \vee \ldots \vee \varphi_n)$ and $\prod_{i=1}^{n} \varphi_i =_{def} (\varphi_1 \wedge \ldots \wedge \varphi_n)$.)

By standard logic, $\exists y(Rxy \wedge \sum_{i=1}^{n} \prod_{j=1}^{n_i} \beta_{ij})$ is equivalent to $\sum_{i=1}^{n} \exists y(Rxy \wedge \prod_{j=1}^{n_i} \beta_{ij})$. So it suffices to consider the members of this disjunction. If none of the β_{ij}'s have a free variable y, then $\exists y(Rxy \wedge \prod_{j=1}^{n_i} \beta_{ij})$ is equivalent to $\exists y(Rxy \wedge (Py \vee \neg Py)) \wedge \prod_{j=1}^{n_i} \beta_{ij}$, for an arbitrary unary predicate constant P. This is a Boolean combination of *M*-formulas of the required kind. Otherwise, let β_i^1 be the conjuction of the β_{ij}'s with y as their free variable and let β_i^2 be the conjunction of the remainder. Then $\exists\, y(Rxy \wedge \prod_{j=1}^{n_i} \beta_{ij})$ is equivalent to $\exists\, y(Rxy \wedge \beta_i^1) \wedge \beta_i^2$, again a Boolean combination of M-formulas of the required kind. QED.

3.5 *Corollary.* Any m-formula with one free variable is equivalent to an M-formula.

Proof. A Boolean combination of M-formulas with the same free variable is itself an M-formula. QED.

m-formulas are obvious generalizations of modal formulas. Note that each m-formula α is, by 3.4, equivalent to a conjunction $\prod_{i=1}^{n} \alpha_i$ of disjunctions $\alpha_i = \sum_{j=1}^{n_i} \alpha_{ij}$ of M-formulas α_{ij}. In many cases, one is only interested in universal closures of such formulas (cf. the definition of "$\mathbf{M} \models \varphi$"). Now, the universal closure $\bar{\alpha}$ of α is equivalent to $\prod_{i=1}^{n} \bar{\alpha}_i$, and, since α_i may be taken to consist of M-formulas α_{ij}, each with a different free variable, $\bar{\alpha}_i$ is equivalent to $\sum_{j=1}^{n_i} \bar{\alpha}_i$. So, in this case, again, (closures of) modal formulas play an important role.

The next result is a semantic characterization of m-formulas (as a subset of all L_1-formulas) in terms of invariance for *generated submodels* and *p-relations*. By 3.5, this yields a characterization of the "modal formulas" in L_1 as well. The relevant definitions are as follows.

For the notion of a "generated submodel", cf. definition 2.10.

3.6 *Definition.* An L_1-formula φ with the free variables $x_1, \ldots, x_n$ is *invariant for generated submodels* if, for all models $\mathbf{M}_1$ and $\mathbf{M}_2$ such that $\mathbf{M} \underset{\rightarrow}{\subset} \mathbf{M}_2$, and for all $w_1, \ldots, w_n \in W_1$, $\mathbf{M}_1 \models \varphi[w_1, \ldots, w_n]$ iff $\mathbf{M}_2 \models \varphi[w_1, \ldots, w_n]$.

3.7 *Definition.* C is a *p-relation* between $\mathbf{M}_1 = \langle W_1, R_1, V_1\rangle$ and $\mathbf{M}_2 = \langle W_2, R_2, V_2\rangle$ if the following four conditions are satisfied:

(i) the domain of C is W_1 and the range of C is W_2
(ii) for each $w \in W_1$ and $v \in W_2$ such that Cwv, and for each unary

predicate constant P, $w \in V_1(P)$ iff $v \in V_2(p)$

(iii) for each w, $w' \in W_1$ and $v \in W_2$ such that R_1ww' and Cwv there exists a $v' \in W_2$ with R_2vv' and $Cw'v'$

(iv) for each v, $v' \in W_2$ and $w \in W_1$ such that R_2vv' and Cwv there exists a $w' \in W_1$ with R_1ww' and $Cw'v'$.

3.8 *Definition.* An L_1-formula φ with the free variables $x_1, \ldots, x_n$ is *invariant for p-relations* if, for all models $\mathbf{M}_1$ and $\mathbf{M}_2$, all p-relations C between $\mathbf{M}_1$ and $\mathbf{M}_2$, and all $w_1, \ldots, w_n \in W_1$, $w'_1, \ldots w'_n \in W_2$ such that $Cw_1w'_1, \ldots, Cw_nw'_n$, $\mathbf{M}_1 \models \varphi[w_1, \ldots, w_n]$ iff $\mathbf{M}_2 \models \varphi[w'_1, \ldots, w'_n]$.

These concepts are of interest only for formulas with free variables. An L_1-*sentence* invariant for generated submodels is either universally valid or a contradiction, as is easily seen using the methods of chapter 2.

3.9 *Theorem.* An L_1-formula φ containing at least one free variable is equivalent to an m-formula iff it is invariant for generated submodels and p-relations.

Proof. One direction is easy. Each m-formula is invariant for generated submodels and p-relations; as a simple formula induction will show.

On the other hand, let φ have this property and let the free variables of φ be $x_1, \ldots, x_n$. Define $\mathrm{m}(\varphi) = \{\psi \mid \psi$ is an m-formula, $\varphi \models \psi$, and the free variables of ψ are among $x_1, \ldots, x_n\}$. We will show that $\mathrm{m}(\varphi) \models \varphi$. By the compactness theorem, this implies $\psi \models \varphi$ for some $\psi \in \mathrm{m}(\varphi)$, whence, clearly, $\models \varphi \leftrightarrow \psi$. Since this proof uses a construction which recurs at various places in chapter 15, it will be given in quite some detail.

Let $\mathbf{M}_1 \models \mathrm{m}(\varphi)[w_1, \ldots, w_n]$. Introduce individual constants $\mathbf{w}_1, \ldots, \mathbf{w}_n$. (The notation $\mathbf{w}$ is consistently used to introduce a unique individual constant for an object w.) Adding $\mathbf{w}_1, \ldots, \mathbf{w}_n$ to L_1 gives a language L_{11}. $\mathbf{M}_1$ is then expanded to an L_{11}-model $\mathbf{M}_{11}$ by interpreting $\mathbf{w}_1$ as $w_1, \ldots, \mathbf{w}_n$ as w_n. Let φ^* be the result of substituting $\mathbf{w}_1$ for $x_1, \ldots, \mathbf{w}_n$ for x_n in φ.

Define $m(L_{11})$ to be the class of those sentences (!) of L_{11} that are obtained by starting with atomic formulas of the form Px or Pc and applying $\neg$, $\wedge$, $\forall y(Rxy \rightarrow$ or $\forall y(Rcy \rightarrow$, where x and y are distinct individual variables and c is an arbitrary individual constant of L_{11}. (m-formulas always had at least one free variable, but this relaxation of the definition generates sentences as well.)

Each finite subset of $\{\varphi^*\} \cup \{\psi \mid \psi \in m(L_{11})$ and $\mathbf{M}_{11} \models \psi\}$ has a model. For, suppose otherwise. Then, for some $\psi_1, \ldots, \psi_k$ as described, $\varphi^* \models \neg(\psi_1 \wedge \ldots \wedge \psi_k)$; but, since $\mathbf{M}_1 \models \mathrm{m}(\varphi)[w_1, \ldots, w_n]$, it follows that $\mathbf{M}_{11} \models \neg(\psi_1 \wedge \ldots \wedge \psi_k)$, contradicting $\mathbf{M}_{11} \models \psi_1 \wedge \ldots \wedge \psi_k$. So, there exists a model $\mathbf{N}_{11}$ for the whole set. $\mathbf{N}_{11}$ is an L_{11}-model satisfying the following two conditions,

(i) $\mathbf{N}_{11} \models \varphi^*$
(ii) $\mathbf{N}_{11} - m(L_{11}) - \mathbf{M}_{11}$,

where (ii) is short for "for each $\varphi \in m(L_{11})$, $\mathbf{N}_{11} \models \varphi$ iff $\mathbf{M}_{11} \models \varphi$".

For each c and w such that c is an individual constant in L_{11}, w is an element of the domain of $\mathbf{N}_{11}$, and $\mathbf{N}_{11} \models Rcx[w]$, add a new constant k_{cw} to L_{11} to obtain L_2. Then expand $\mathbf{N}_{11}$ to an L_2-model $\mathbf{N}_2$ by interpreting each k_{cw} as w. $m(L_2)$ is defined in the obvious way.

Each finite subset of $\{\psi \mid \psi \in m(L_2)$ and $\mathbf{N}_2 \models \psi\} \cup \{Rck_{cw} \mid \mathbf{N}_2 \models Rck_{cw}\}$ has a model which is an expansion of $\mathbf{M}_{11}$. To prove this, consider $\psi_1, \ldots, \psi_k$ as described, together with $Rc_1k_{c_1w_1}, \ldots, Rc_lk_{c_lw_l}$. Add Rck_{cw} for each k_{cw} occurring in $\psi_1 \wedge \ldots \wedge \psi_k$ which is not among $k_{c_1w_1}, \ldots, k_{c_lw_l}$; say for $k_{c'_1w'_1}, \ldots, k_{c'_sw'_s}$. Then take distinct variables $x_1, \ldots, x_l, y_1, \ldots, y_s$ not occurring in $\psi_1 \wedge \ldots \wedge \psi_k$ and substitute them for $k_{c_1w_1}, \ldots, k_{c_lw_l}, k_{c'_1w'_1}, \ldots, k_{c'_sw'_s}$, respectively, to obtain $(\psi_1 \wedge \ldots \wedge \psi_k)'$. Then $\mathbf{N}_{11} \models \exists x_1(Rc_1x_1 \wedge \ldots \wedge \exists x_l(Rc_lx_l \wedge \exists y_1(Rc'_1y_1 \wedge \ldots \wedge \exists y_s(Rc'_sy_s \wedge (\psi_1 \wedge \ldots \wedge \psi_k)' \ldots)$. This sentence is in $m(L_{11})$ and, therefore, it also holds in $\mathbf{M}_{11}$, since $\mathbf{N}_{11} - m(L_{11}) - \mathbf{M}_{11}$. It is now clear how $\mathbf{M}_{11}$ can be expanded to a model for $\{\psi_1, \ldots, \psi_k, Rc_1k_{c_1w_1}, \ldots, Rc_lk_{c_lw_l}\}$.

Using a well-known model-theoretic argument it follows that the above set has a model $\mathbf{M}_2$ satisfying the following conditions:

(i) $\mathbf{M}_{11} \prec_{L_{11}} \mathbf{M}_2$ (i.e., $\mathbf{M}_{11}$ is an L_{11}-*elementary submodel* of $\mathbf{M}_2$)
(ii) $\mathbf{N}_2 - m(L_2) - \mathbf{M}_2$,

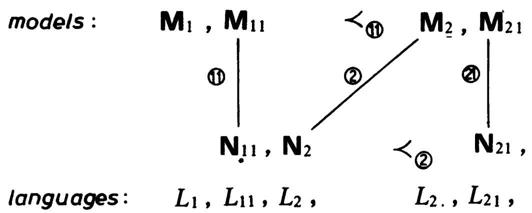
where (ii) has the obvious meaning. This situation may be pictured as:

$$\begin{tikzcd}[arrows=dash] \textit{models}: & \mathbf{M}_1, & \mathbf{M}_{11} \arrow[d, "11"'] & \prec_{11} & \mathbf{M}_2, \arrow[dl, "2"'] \\ & & \mathbf{N}_{11}, & \mathbf{N}_2 & \end{tikzcd}$$

languages : L_1, L_{11}, L_2, L_2, This construction is repeated, but now starting from $\mathbf{M}_2$. For each c and w such that c is a constant in L_2, w is an element in the domain of $\mathbf{M}_2$ and $\mathbf{M}_2 \models Rcx[w]$, add a new constant k_{cw} to L_2 to obtain L_{21}. $\mathbf{M}_2$ is then expanded to an L_{21}-model $\mathbf{M}_{21}$ by interpreting each k_{cw} as w. Using an argument similar to the one given above, one sees that each finite subset of $\{\psi \mid \psi \in m(L_{21}) \text{ and } \mathbf{M}_{21} \models \psi\} \cup \{Rck_{cw} \mid k_{cw} \in L_{21}\text{-}L_2 \text{ and } \mathbf{M}_{21} \models Rck_{cw}\}$ has a model which is an expansion of $\mathbf{N}_2$. Therefore, this set has a model $\mathbf{N}_{21}$ satisfying the following two conditions,

(i) $\mathbf{N}_2 \prec_{L_2} \mathbf{N}_{21}$
(ii) $\mathbf{N}_{21} - m(L_{21}) - \mathbf{M}_{21}$.

In the picture this leads to:



Iterating this construction yields two elementary chains $\mathbf{M}_1, \mathbf{M}_2, \ldots$ and $\mathbf{N}_{11}, \mathbf{N}_{21}, \ldots$ with limits $\mathbf{M}$ and $\mathbf{N}$, respectively. Moreover, the interpreted constants of the total language created in the process form two special submodels on both sides — with zigzag connections between R-successive objects built in step-by-step, through the above Rck_{cw}-formulas. The required conclusion then follows from the assumptions on φ and the fundamental theorem on elementary chains. First, since $\mathbf{N}_{11} \models \varphi^*$, $\mathbf{N} \models \varphi^*$. The submodel $\mathbf{N}_c$ of $\mathbf{N}$ generated by the constants in $\bigcup_n L_n$ is a generated submodel of $\mathbf{N}$ and, therefore, $\mathbf{N}_c \models \varphi^*$, by the invariance of φ for generated submodels. The following now defines a p-relation C between $\mathbf{N}_c$ and the generated submodel $\mathbf{M}_c$ of $\mathbf{M}$ generated by the constants of $\bigcup_n L_n$. Define Cwv to hold if, for some constant $c \in \bigcup_n L_n$, $w = c^{\mathbf{M}}$ and $v = c^{\mathbf{N}}$. The construction of the chains guarantees that C satisfies the four properties required. By the invariance of φ for p-relations, $\mathbf{M}_c \models \varphi^*$, and, using the invariance of φ for generated submodels once more, $\mathbf{M} \models \varphi^*$. This implies that $\mathbf{M}_{11} \models \varphi^*$, so $\mathbf{M}_1 \models \varphi[w_1, \ldots, w_n]$. QED.

The use of constants k_{cw}, rather than just w, in this proof serves to avoid the following complication. Let c_1 and c_2 be constants of L_{11} and let $\mathbf{N}_2 \models Rc_1x[w]$ and $\mathbf{N}_2 \models Rc_2x[w]$. $\{Rc_1w, Rc_2w\}$ need not have a model which is an expansion of $\mathbf{M}_{11}$. (The method used only leads to the L_{11}-sentence $\exists x_1(Rc_1x_1 \wedge Rc_2x_1)$; but this is not a sentence in $m(L_{11})$, and therefore need not be true in $\mathbf{M}_{11}$.) Using $k_{c_{1w}}$ and $k_{c_{2w}}$ leads to the $m(L_{11})$-sentence $\exists x_1(Rc_1x_1 \wedge \exists x_2Rc_2x_2)$, in which the information about c_1 and c_2's having a common R-successor is lost.

This concludes the part on modal formulas as L_1-formulas. We have now arrived at the main point of interest, viz. modal formulas as L_2-formulas, i.e., as means of expressing (second-order) properties of the alternative relation. The above correspondence connecting a modal formula φ with the L_2-formula $\forall P_1 \ldots \forall P_nST(\varphi)$ is rather uninformative, however. In many cases, simpler equivalents may be found, often even *first-order* properties (i.e., L_0-formulas). A few examples will give an impression of this.

For all frames $\mathbf{F}$ $(= \langle W, R\rangle)$ and all $w \in W$,

(1) $\mathbf{F} \models L\perp[w]$ iff $\mathbf{F} \models \neg\exists yRxy[w]$
(2) $\mathbf{F} \models M\top[w]$ iff $\mathbf{F} \models \exists yRxy[w]$
(3) $\mathbf{F} \models ML\perp[w]$ iff $\mathbf{F} \models \exists y(Rxy \wedge \neg\exists zRyz)[w]$

(In these first examples, valuations play no role at all.)

(4) $\mathbf{F} \models Lp \rightarrow p[w]$ iff $\mathbf{F} \models Rxx[w]$
(5) $\mathbf{F} \models Lp \rightarrow LLp[w]$ iff $\mathbf{F} \models \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))[w]$
(6) $\mathbf{F} \models MLp \rightarrow p[w]$ iff $\mathbf{F} \models \forall y(Rxy \rightarrow Ryx)[w]$.

And, finally comes an example showing that identity may be needed as well:

(7) $\mathbf{F} \models p \rightarrow Lp[w]$ iff $\mathbf{F} \models \forall y(Rxy \rightarrow y = x)[w]$.

Proofs of these equivalences will be found in chapter 9.

Some readers will be used to equivalences of the form

(4)′ $\mathbf{F} \models Lp \rightarrow p$ iff $\mathbf{F} \models \forall xRxx$.

We prefer the "parametrized" form given here, because it is more natural (as will appear in chapter 9) and because it is more informative, being stronger than the second type of formulation (as will appear in chapter 7).

To give a first idea of how such equivalences are proven, $CF = L((Lp \wedge p) \rightarrow q) \vee L(Lq \rightarrow p)$ will serve as a not too trivial example:

(8) $\mathbf{F} \models CF[w]$ iff $\mathbf{F} \models \forall y(Rxy \rightarrow \forall z(Rxz \rightarrow (Ryz \vee Rzy \vee z = y)))[w]$.

(I.e., CF defines a kind of *connectedness*.)

Proof of (8). From right to left: Suppose the given relational property holds at w. Let V be any valuation on $\mathbf{F}$. If $\langle \mathbf{F}, V\rangle \models L((Lp \wedge p) \rightarrow q)[w]$, there is nothing to be proven, so suppose $\langle \mathbf{F}, V\rangle \not\models L((Lp \wedge p) \rightarrow q)[w]$. This means that, for some $v \in W$ with Rwv, $\langle \mathbf{F}, V\rangle \not\models (Lp \wedge p) \rightarrow q[v]$, i.e., $\langle \mathbf{F}, V\rangle \models Lp \wedge p[v]$, but $\langle \mathbf{F}, V\rangle \not\models q[v]$. Now, it is to be shown that $\langle \mathbf{F}, V\rangle \models L(Lq \rightarrow p)[w]$; so let u be any R-successor of w such that $\langle \mathbf{F}, V\rangle \models Lq[u]$. It suffices to show that $\langle \mathbf{F}, V\rangle \models p[u]$. There are at most three possibilities: either Rvu (and $\langle \mathbf{F}, V\rangle \models Lp[v]$ yields the desired conclusion), or Ruv (which is impossible, since $\langle \mathbf{F}, V\rangle \models Lq[u]$ and $\langle \mathbf{F}, V\rangle \not\models q[v]$), or $u = v$ (and $\langle \mathbf{F}, V\rangle \models p[v]$ yields the desired conclusion).

From left to right: suppose the given relational property fails at w. This means that there are v, u such that Rwv, Rwu, $\sim Rvu$, $\sim Ruv$ and $u \neq v$. Define a valuation V by setting $V(p) = \{v\} \cup \{x \in W \mid Rvx\}$ and $V(q) = \{x \in W \mid Rux\}$. An easy calculation shows that $\langle \mathbf{F}, V\rangle \not\models (Lp \wedge p) \rightarrow q[v]$ and $\langle \mathbf{F}, V\rangle \not\models Lq \rightarrow p[u]$, whence $\langle \mathbf{F}, V\rangle \not\models CF[w]$. **QED.**

An important example of a modal principle which is *not* definable by means of an L_0-formula is "Löb's Formula" $L(Lp \rightarrow p) \rightarrow Lp$ (LF; cf. chapter 1):

(9) $\mathbf{F} \models LF[w]$ iff (i) $\mathbf{F} \models \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))[w]$ and (ii) there is no $f : \mathbb{N} \rightarrow W$ such that $f(0) = w$ and $Rf(n)f(n+1)$ for each $n \in \mathbb{N}$.

(Consequently, $\mathbf{F} \models LF$ iff R is *transitive* and the converse relation of R is *well-founded*.)

Proof of (9). From left to right: That LF implies transitivity is seen most elegantly by considering LF as the L_2-formula $\forall P(\forall y(Rxy \rightarrow (\forall z(Ryz \rightarrow Pz) \rightarrow Py)) \rightarrow \forall y(Rxy \rightarrow Py))$ and substituting the L_0-formula $Rxu \wedge \forall v(Ruv \rightarrow Rxv)$ for subformulas of the form Pu. That LF implies well-foundedness of the converse of R is seen by contraposition: Suppose an f as forbidden exists. Define $V(p) = W - \{f(n) \mid n \in \mathbb{N}\}$. For no $f(n)$,

$\langle \mathbf{F}, V\rangle \models Lp[f(n)]$, and, for all v's which are not $f(n)$'s, $\langle \mathbf{F}, V\rangle \models p[v]$. Therefore, $\langle \mathbf{F}, V\rangle \models L(Lp \to p)[f(0)]$, and $\langle \mathbf{F}, V\rangle \not\models Lp[f(0)]$: LF has been falsified.

From right to left: Assume that, for some V and w, $\langle \mathbf{F}, V\rangle \not\models LF[w]$ (i.e., $\langle \mathbf{F}, V\rangle \models L(Lp \to p)[w]$ and $\langle \mathbf{F}, V\rangle \not\models Lp[w]$), but (i) holds. An f as forbidden may then be constructed as follows. Set $f(0) = w$. Now, suppose that $f(n)$ has been defined already such that $Rf(0)f(1), \ldots, Rf(n\text{-}1)f(n)$ and $\langle \mathbf{F}, V\rangle \not\models Lp[f(n)]$. This means that, for some $v \in W$ with $Rf(n)v$, $\langle \mathbf{F}, V\rangle \not\models p[v]$. Because of (i), $Rf(0)f(n)$ holds, and, therefore, also $Rf(0)v$. Now, $\langle \mathbf{F}, V\rangle \models L(Lp \to p)[f(0)]$, whence $\langle \mathbf{F}, V\rangle \models Lp \to p[v]$, so $\langle \mathbf{F}, V\rangle \not\models Lp[v]$. Set $v = f(n+1)$. **QED.**

As will be shown below, well-foundedness of the converse relation of R alone is not modally definable.

A formula which behaves like LF in many respects is the following variant of "Dummett's Axiom": $Dum = L(L(p \to Lp) \to p) \to p$. The relevant equivalence is stated here without proof:

(10) $\mathbf{F} \models Dum[w]$ iff (i) $\mathbf{F} \models Rxx[w]$ and
(ii) $\mathbf{F} \models \forall y(Rxy \to \forall z(Ryz \to Rxz))[w]$
and (iii) there is no $f{:}IN \to W$ such that $f(0)=w$ and for each $n \in IN$, $Rf(n)f(+1)$ and $f(n) \neq f(n+1)$.

Now that these examples have been given, we are ready for the main notion:

3.10 *Definition.* For any modal formula φ and any L_0-formula α with one free variable,

$E(\varphi, \alpha)$ iff $\forall \mathbf{F}\ (= \langle W, R\rangle)\ \forall w \in W(\mathbf{F} \models \varphi[w] \Leftrightarrow \mathbf{F} \models \alpha[w])$.

For any modal formula φ and any L_0-sentence α,

$\bar{E}(\varphi, \alpha)$ iff $\forall \mathbf{F}(\mathbf{F} \models \varphi \Leftrightarrow \mathbf{F} \models \alpha)$
$M1 = \{\varphi \mid \text{for some } \alpha, E(\varphi, \alpha)\}$
$\bar{M}1 = \{\varphi \mid \text{for some } \alpha, \bar{E}(\varphi, \alpha)\}$ $(= \{\varphi \mid FR(\varphi)$ is L_0-*elementary*$\})$
$P1 = \{\alpha \mid \text{for some } \varphi, E(\varphi, \alpha)\}$
$\bar{P}1 = \{\alpha \mid \text{for some } \varphi, \bar{E}(\varphi, \alpha)\}$.

E is the relation of *local* equivalence between a modal formula and an L_0-formula, and $\bar{E}$ is the relation of *global* equivalence. If $E(\varphi, \alpha)$, where α has the one free variable x, then $\bar{E}(\varphi, \forall x\alpha)$. It follows that $M1 \subseteq \bar{M}1$ and $P1 \subseteq \bar{P}1$. The first inclusion is proper, as will be shown in chapter 7. For the second one, this is an open (but less interesting) question.

In the course of chapter 2, several L_0-sentences have been shown to be outside of $\bar{P}1$, viz. $\exists x \exists yRxy$, $\forall x\forall yRxy$, $\forall x\neg Rxx$ and $\forall x \exists y(Rxy \wedge Ryy)$. This information will have to do until chapter 14, when the study of $P1$ and $\bar{P}1$ is taken up. In the meantime we will be concerned with $M1$ and $\bar{M}1$ only.

There are two ways in which the above definition may be generalized. First, one may consider the class of modal formulas φ such that φ is equivalent to some *set* of L_0-formulas. There is no gain in this direction, however. To consider the most interesting case, suppose $FR(\varphi) = FR(\Sigma)$ for some set Σ of L_0-sentences (i.e., $FR(\Sigma)$ is L_0-Δ-*elementary*). This implies that $\Sigma \models \forall P_1 \ldots \forall P_n\forall xST(\varphi)$, whence also $\Sigma \models \forall xST(\varphi)$. Now all sentences involved here are L_1-sentences, whence, by the compactness theorem for L_1, some finite $\Sigma_0 \subseteq \Sigma$ exists with $\Sigma_0 \models \forall xST(\varphi)$ (and hence $\Sigma_0 \models \forall P_1 \ldots \forall P_n\forall xST(\varphi)$). It is easily seen that the conjunction σ of all formulas in Σ_0 defines $FR(\varphi)$ as well. In fact, we have only hit upon a special case of the following general result:

For any universal second-order sentence φ, if φ defines a Δ-elementary class of structures, then it defines an elementary class of structures.

Secondly, one may consider *sets* of modal formulas instead of single ones. E.g., let $\Sigma \in \bar{M}1$ if $FR(\Sigma)$ is L_0-Δ-elementary. (The definition of $M1$ is extended analogously.) There is no reduction like above now, witness the following example from [6].

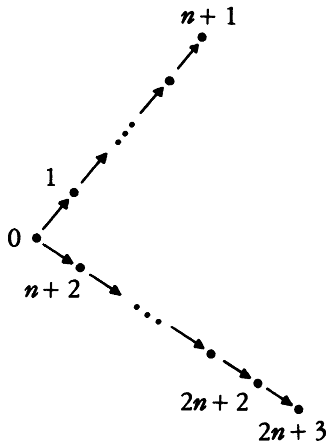
Let $\Sigma = \{M^iL\bot \rightarrow L^iL\bot \mid i \geqslant 1\}$. Each $M^iL\bot \rightarrow L^iL\bot$ defines an L_0-property, viz. $\exists x_1(Rxx_1 \wedge \ldots \wedge \exists x_i(Rx_{i-1}x_i \wedge \neg \exists yRx_iy)) \rightarrow \forall x_1(Rxx_1 \rightarrow \ldots \rightarrow \forall x_i(Rx_{i-1}x_i \rightarrow \neg \exists yRx_iy)\ldots)$. Hence, Σ defines an L_0-Δ-elementary class of frames. But this class is not L_0-elementary. For, suppose it were; say the L_0-sentence σ defined it. Then, by compactness, σ would follow from some finite number of formulas of the form $M^iL\bot \rightarrow L^iL\bot$ and, therefore, for some n, it would follow from $\{M^iL\bot \rightarrow L^iL\bot \mid 1 \leqslant i \leqslant n\}$.

Consequently, this set would imply (amongst others) $M^{n+1}L\bot \rightarrow L^{n+1}L\bot$, which cannot be the case. To refute it, consider the frame $\mathbf{F} = \langle W, R\rangle$ with

$W = \{0, ..., 2n+3\}$

and

$R = \{\langle i, i+1\rangle \mid 0 \leqslant i \leqslant n\} \cup \{\langle 0, n+2\rangle\} \cup \{\langle n+i, n+i+1\rangle \mid 2 \leqslant i \leqslant n+2\}$:



$\mathbf{F} \models M^{n+1}L\bot[0]$, but $\mathbf{F} \not\models L^{n+1}L\bot[0]$, whence $M^{n+1}L\bot \rightarrow L^{n+1}L\bot$ is falsified in $\mathbf{F}$. But, all formulas $M^iL\bot \rightarrow L^iL\bot$ $(1 \leqslant i \leqslant n)$ hold in $\mathbf{F}$, as is easily checked. QED.

Using sets of modal formulas instead of single modal formulas has certain advantages. E.g., the difference between $M1$ and $\bar{M}1$ becomes negligible:

3.11 *Lemma.* For any set Σ of modal formulas which is closed under prefixing of L, $\Sigma \in M1$ iff $\Sigma \in \bar{M}1$.

Proof. If $\Sigma \in M1$, then, for some set Δ of L_0-formulas all having the same free variable (say x), $\mathbf{F} \models \Sigma[w]$ iff $\mathbf{F} \models \Delta[w]$, for all $\mathbf{F}$ and w. Therefore, for all $\mathbf{F}$, $\mathbf{F} \models \Sigma$ iff $\mathbf{F} \models \{\forall x\delta \mid \delta \in \Delta\}$, and so $\Sigma \in \bar{M}1$.

If $\Sigma \in \bar{M}1$, then, for some set Δ of L_0-sentences, $\mathsf{F} \models \Sigma$ iff $\mathsf{F} \models \Delta$, for all F. By the preservation results of chapter 15, Δ's being preserved under generated subframes and disjoint unions (because Σ is) implies that Δ is axiomatized by its logical consequences of the form $\forall x\delta$, where $\delta = \delta(x)$ is a *restricted* formula (i.e., a formula in which the only occurrences of quantifiers are of the forms $\forall y(Rzy \rightarrow$ or $\exists y(Rzy \wedge$, with y distinct from z). Note that restricted formulas are *invariant for generated subframes* in the following sense: if $\mathsf{F}_1 \underset{\rightarrow}{\subseteq} \mathsf{F}_2$ and $w \in W_1$, then $\mathsf{F}_1 \models \delta[w]$ iff $\mathsf{F}_2 \models \delta[w]$. (This may be shown using a formula induction similar to the one establishing lemma 2.11.) Now, it will be shown that, for all F and w, $\mathsf{F} \models \Sigma[w]$ iff $\mathsf{F} \models \Delta'[w]$; where $\Delta' = \{\delta \mid \delta$ is restricted and $\Delta \models \forall x\delta\}$.

If $\mathsf{F} \models \Sigma[w]$, then, by 2.12, $TC(\mathsf{F}, w) \models \Sigma[w]$ and, by Σ's being closed under prefixing of L, $TC(\mathsf{F}, w) \models \Sigma$. The assumption about Σ then yields that $TC(\mathsf{F}, w) \models \Delta$, whence $TC(\mathsf{F}, w) \models \Delta'[w]$. By the invariance of the formulas in Δ' for generated subframes, $\mathsf{F} \models \Delta'[w]$.

Conversely, let $\mathsf{F} \models \Delta'[w]$. As above, it follows that $TC(\mathsf{F}, w) \models \Delta'[w]$. But this implies that $TC(\mathsf{F}, w) \models \Delta$ (and, once this has been established, it suffices to note that $TC(\mathsf{F}, w) \models \Sigma$, $TC(\mathsf{F}, w) \models \Sigma[w]$ and hence $\mathsf{F} \models \Sigma[w]$). To see that $TC(\mathsf{F}, w) \models \Delta$, it suffices, by the above remark about Δ, to prove that, for any $\delta \in \Delta'$, and any v in the domain of $TC(\mathsf{F}, w)$, $TC(\mathsf{F}, w) \models \delta[v]$. So, let v be connected with w by an R-chain of length n. Since $\Delta \models \forall x\delta$, it also holds that

$$\Delta \models \forall x \forall y_1(Rxy_1 \rightarrow \ldots \rightarrow \forall y_n(Ry_{n-1}y_n \rightarrow \delta(y_n))\ldots),$$

which will be abbreviated as $\forall x \forall y(R^n xy \rightarrow \delta(y))$. This formula is restricted and hence belongs to Δ'. So, $TC(\mathsf{F}, w) \models \forall y(R^n xy \rightarrow \delta(y))[w]$, and the required conclusion $TC(\mathsf{F}, w) \models \delta[v]$ follows easily. QED.

In chapter 5, *modal logics* will be defined as sets of modal formulas which are (amongst others) closed under prefixing of L. It follows from 3.11 that, for modal logics, there is no difference between local and global first-order definability.

Generally speaking, $\bar{M}1$ will turn out to be a more natural object for semantic study (cf. chapter 8), but $M1$ for syntactic study (cf. chapter 9). Recall also the remark made at the end of chapter 2: if the basic structures are taken to be of the form $\langle \mathsf{F}, w \rangle$ with some distinguished world w, then $M1$ is the more natural class anyway.

When the class of modal formulas is viewed as a subclass of all L_2-formulas, one becomes interested in the same kind of question which inspired 3.4 and 3.9. For an example of a syntactic result in this area, the reader is referred to 2.9, where it was shown how $\rightarrow$ and M suffice as logical primitives. Semantic results are hard to obtain, however. It is tempting (in view of the notions introduced in chapter 2) to conjecture that any L_2-sentence which is universal with respect to its second-order quantifiers is equivalent to a modal formula iff it is preserved under generated subframes, disjoint unions and p-morphisms, while its negation is preserved under ultrafilter extensions. But, the following sentence defining well-foundedness of the converse relation of R is a counterexample, as will be seen below:

$$\forall P(\forall x(\forall y(Rxy \rightarrow Py) \rightarrow Px) \rightarrow \forall xPx).$$

(Note the curious fact that well-foundedness of the converse relation of R together with transitivity of R *is* modally definable, viz. by Löb's Formula $L(Lp \rightarrow p) \rightarrow Lp$.) The above sentence is preserved under subframes, whence it is preserved under generated subframes and its negation is preserved under ultrafilter extensions (use 2.27). That it is preserved under disjoint unions and p-morphisms is not hard to check either. Now, to see that it is not modally definable, suppose that the set Σ of modal formulas defined this sentence. Then, consider the frame $\langle IN, R\rangle$, where $R = \{\langle n, n+1\rangle \mid n \in IN\}$. Clearly, the converse relation of R is not well-founded; whence, for some valuation V, some $n \in IN$ and some $\sigma \in \Sigma$, $\langle IN, R, V\rangle \models \neg\, \sigma[n]$. By lemma 2.23, some finite submodel of this model falsifies σ as well. But, the converse relation of R is well-founded in such a finite model: a contradiction.

The difficulty in finding preservation results in this area is further illustrated by the fact that the above sentence, although preserved under subframes, is not definable by means of a universal second-order sentence of the form:

$$\forall P_1 \dots \forall P_m \forall x_1 \dots \forall x_n \varphi,$$

where $\varphi = \varphi(P_1, \dots, P_m, R, =)$ is a first-order formula in which no quantifiers occur. (I.e., Tarski's well-known first-order preservation result involving substructures and universal sentences fails here.) The reason is that well-foundedness is not definable by means of L_0-sentences, but universal second-order sentences of the described kind are:

3.12 *Lemma.* Any second-order sentence of the form $\forall P_1 \ldots \forall P_m \forall x_1 \ldots \forall x_n \varphi$, where φ is a first-order formula constructed from L_0-formulas and atomic formulas of the forms $P_i y_1 \ldots y_{n_i}$ $(1 \leqslant i \leqslant m)$ using Boolean operators only, is preserved under ultraproducts.

Proof. Let $\mathbf{F}_i \models \forall P_1 \ldots \forall P_m \forall x_1 \ldots \forall x_n \varphi$, for each i in some index set I. Let U be any ultrafilter on I. It will be proven that $\mathbf{F} = \Pi_U \mathbf{F}_i \models \forall P_1 \ldots \forall P_m \forall x_1 \ldots \forall x_n \varphi$. To see this, let $Y_1, \ldots, Y_m$ be arbitrary predicates on $\mathbf{F}$ such that, if P_i is n_i-ary, then Y_i is n_i-ary $(1 \leqslant i \leqslant m)$. Moreover, let $f^1_U, \ldots, f^n_U$ be arbitrary elements of W. It is to be shown that $\langle \mathbf{F}, Y_1, \ldots, Y_m \rangle \models \varphi[f^1_U, \ldots, f^n_U]$. Now, φ is a Boolean combination of L_0-formulas $\alpha_1, \ldots, \alpha_s$ and atomic formulas of the form $P_i y_1 \ldots y_{ni}$, say $\beta_1, \ldots, \beta_t$. We shall find a factor $\mathbf{F}_i$ such that

(i) for each $j (1 \leqslant j \leqslant s)$, $\mathbf{F}_i \models \alpha_j[f^1(i), \ldots, f^n(i)]$ iff $\mathbf{F} \models \alpha_j[f^1_U, \ldots, f^n_U]$,

and also predicates $Y^i_1, \ldots, Y^i_m$ on $\mathbf{F}_i$ such that

(ii) for each k $(1 \leqslant k \leqslant t)$, $\langle \mathbf{F}_i, Y^i_1, \ldots, Y^i_m \rangle \models \beta_k[f^1(i), \ldots, f^n(i)]$ iff $\langle \mathbf{F}, Y_1, \ldots, Y_m \rangle \models \beta_k[f^1_U, \ldots, f^n_U]$.

Since $\mathbf{F}_i \models \forall P_1 \ldots \forall P_m \forall x_1 \ldots \forall x_n \varphi$, it will hold that

(iii) $\langle \mathbf{F}, Y^i_1, \ldots, Y^i_m \rangle \models \varphi[f^1(i), \ldots, f^n(i)]$.

(i), (ii) and (iii) combined yield that $\langle \mathbf{F}, Y_1, \ldots, Y_m \rangle \models \varphi[f^1_U, \ldots, f^n_U]$.

Now, to find $\mathbf{F}_i$. Because of the Theorem of Łoś, $\mathbf{F} \models \alpha_j[f^1_U, \ldots, f^n_U]$ iff $\{i \in I \mid \mathbf{F}_i \models \alpha_j[f^1(i), \ldots, f^n(i)]\} \in U$, and the same holds for $\neg\alpha_j$. Taking the appropriate intersection of s sets in the ultrafilter yields a set A in U of indices i satisfying (i). A similar consideration yields a set B in U of indices i such that, for all j, k $(1 \leqslant j, k \leqslant n)$ $f^j(i) = f^k(i)$ iff $f^j_U = f^k_U$. Clearly, $A \cap B$ is in U, whence, in particular, it is non-empty. Now any $i \in A \cap B$ will do. It clearly satisfies (i) and, moreover, predicates $Y^i_1, \ldots, Y^i_m$ as required may be defined by setting: Y^i_j is the set of n_j-tuples $\langle f^{k_1}(i), \ldots, f^{k_{n_j}}(i) \rangle$ such that $Y_j(f^{k_1}_U, \ldots, f^{k_{n_j}}_U)$. By the fact that $i \in B$, it can easily be seen that, if $Y_j(f^{k_1}_U, \ldots, f^{k_{n_j}}_U)$ does *not* hold, then $Y^i_j(f^{k_1}(i), \ldots, f^{k_{n_j}}(i))$ will not either. Hence (ii) is satisfied. QED.

3.13 *Corollary.* Any second-order sentence of the form described in lemma 3.12 is definable by means of a single first-order formula.

Proof. Let φ be such a second-order sentence. Clearly, both φ and $\neg\varphi$ are preserved under isomorphisms. Moreover, $\neg\varphi$ is preserved under ultraproducts, being equivalent to an existential second-order sentence (cf. Chang & Keisler [17], corollary 4.1.14). Lemma 3.12 says that φ is preserved under ultraproducts as well; whence the conditions of Keisler's well-known characterization of elementary classes (cf. Chang & Keisler [17], corollary 6.1.16) are satisfied: φ is definable by means of a single first-order sentence. QED.

The further study of modal formulas as L_2-formulas, and especially their connection with L_0-formulas, will be taken up in chapter 7.

CHAPTER IV

MODAL ALGEBRAS

In this chapter, some basic notions concerning modal algebras and their connection with possible worlds semantics will be presented. For a detailed exposition, cf. R. I. Goldblatt [30].

4.1 *Definition.* A *modal algebra* is an ordered quintuple $\mathbf{A} = \langle A, 1, -, \cap, l \rangle$, where $\langle A, 1, -, \cap \rangle$ is a Boolean algebra and l is a unary operation on A satisfying

(i) $l(x \cap y) = l(x) \cap l(y)$
(ii) $l(1) = 1$.

If Boolean operators are interpreted in the appropriate manner and L is made to correspond to the operation l, then modal formulas may be regarded as polynomials which can be interpreted in these algebras in an obvious way. More precisely, the *polynomial transcription* $\bar{\varphi}$ of a modal formula φ may be defined inductively, according to the clauses:

(i) $\bar{\top} = 1$, $\bar{\bot} = 0$ $(= -1)$
(ii) $\bar{p}$ is some variable associated with p, for each proposition letter p
(iii) $\overline{\neg\varphi} = -\bar{\varphi}$
(iv) $\overline{\varphi \wedge \psi} = \bar{\varphi} \cdot \bar{\psi}$
(v) $\overline{L\varphi} = l(\bar{\varphi})$.

E.g., $\overline{Lp \rightarrow Mp} = \overline{\neg(Lp \wedge L\neg p)}$ becomes $-(l(x) \cdot l(-x))$. Clearly, an assignment of objects in A to the variables may be lifted to a function from such polynomials to A.

The usual algebraic notions (*homomorphism*, *subalgebra* and *direct product*) apply to modal algebras in a straigthforward manner.

Any frame $\mathbf{F}$ ($= \langle W, R\rangle$) induces a modal algebra $\mathbf{F}^{+} = \langle A, 1, -, \cap, l\rangle$, where A is the power set of W, $1 = W$, $-$ is complementation with respect to W, $\cap$ is intersection and l is the set-theoretic operation defined by $l(X) = \{w \in W \mid \forall v \in W, Rwv$ only if $v \in X\}$, for all $X \subseteq W$. Clearly, $Th_{mod}(\mathbf{F})$ consists of just those modal formulas which, when considered as polynomials, receive the value 1 in $\mathbf{F}^{+}$ regardless of the assignment to their variables.

The well-known *Stone Representation* connects modal algebras $\mathbf{A}$ with frames $\mathbf{F}$ ($= \langle W, R\rangle$), where W is the set of all ultrafilters on $\mathbf{A}$ and R is some suitably defined relation. It turns out that the modal theory $Th_{mod}(\mathbf{F})$ of such an $\mathbf{F}$ need not equal the modal theory of $\mathbf{A}$ (although it is contained in it). To get a better correspondence, it is necessary to introduce the following concept.

4.2 *Definition* (S. K. Thomason [79]). A general frame $\langle \mathbf{F}, \mathbb{W}\rangle$ consists of a frame $\mathbf{F}$ ($= \langle W, R\rangle$) and a non-empty set $\mathbb{W}$ of subsets of W closed under intersections, relative complements with respect to W and the set-theoretic operation l defined above. A modal formula φ *holds at* w *in* $\langle \mathbf{F}, \mathbb{W}\rangle$ ($\langle \mathbf{F}, \mathbb{W}\rangle \models \varphi[w]$) if, for all valuations V on $\mathbf{F}$ taking values in $\mathbb{W}$, $\langle \mathbf{F}, V\rangle \models \varphi[w]$. ($\langle \mathbf{F}, \mathbb{W}\rangle \models \Sigma[w]$, $\langle \mathbf{F}, \mathbb{W}\rangle \models \varphi$, $\langle \mathbf{F}, \mathbb{W}\rangle \models \Sigma$ and $Th_{mod}(\langle \mathbf{F}, \mathbb{W}\rangle)$ are then defined in the obvious way.)

Although general frames were invented as set-theoretic representations of modal algebras, a different motivation may be given for their introduction as well. When dealing with axiomatic theories, one often has a formula φ serving as an axiom *schema*, i.e., all substitution instances of φ count as axioms. Now if $\mathbf{F} \models \varphi$ for some frame $\mathbf{F}$, then, indeed, $\mathbf{F} \models [\psi_1/p_1, \ldots, \psi_n/p_n]\varphi$ for any substitution instance $[\psi_1/p_1, \ldots, \psi_n/p_n]\ \varphi$ of φ. (This follows from lemma 2.5.) But, if all instances of φ are true in some model $\mathbf{M}$, then this fact cannot be summed up by stating that $\mathbf{M} \models \varphi$, since truth of φ in $\mathbf{M}$ need not imply truth of φ's substitution instances in $\mathbf{M}$. By turning $\mathbf{M}$ ($= \langle W, R, V\rangle$) into the general frame $\langle \mathbf{F}, \mathbb{W}\rangle = \langle\langle W, R\rangle, \{V(\psi) \mid \psi$ is a modal formula$\}\rangle$, however, it suffices to state that $\langle \mathbf{F}, \mathbb{W}\rangle \models \varphi$. For, truth of a formula in a general frame implies the truth of all its substitution instances.

Note that standard frames may be taken to be general frames where

$\mathbf{W}$ equals the power set of W.

The terminology "general frame" is taken over from Henkin's "general models" which extend the class of standard (second-order) models just as the general frames extend the class of (standard) frames. (Cf. Henkin [33].)

The four fundamental notions concerning frames are also applicable to general frames:

4.3 *Definition.* A general frame $\langle \mathsf{F}_1, \mathbf{W}_1 \rangle$ is a *generated subframe* of a general frame $\langle \mathsf{F}_2, \mathbf{W}_2 \rangle$ (notation: $\langle \mathsf{F}_1, \mathbf{W}_1 \rangle \subseteq \langle \mathsf{F}_2, \mathbf{W}_2 \rangle$) if F_1 is a generated subframe of F_2 and $\mathbf{W}_1 = \{X \cap W_1 \mid X \in \mathbf{W}_2\}$.

The *disjoint union* $\Sigma\{\langle \mathsf{F}_i, \mathbf{W}_i \rangle \mid i \in I\}$ of a set $\{\langle \mathsf{F}_i, \mathbf{W}_i \rangle \mid i \in I\}$ of general frames is the general frame $\langle \mathsf{F}, \mathbf{W} \rangle$, where $\mathsf{F}(= \langle W, R \rangle) = \Sigma\{\mathsf{F}_i \mid i \in I\}$ and $\mathbf{W} = \{X \subseteq W \mid \text{for each } i \in I,\ X \cap W_i \in \mathbf{W}_i\}$.

A function f from a general frame $\langle \mathsf{F}_1, \mathbf{W}_1 \rangle$ onto a general frame $\langle \mathsf{F}_2, \mathbf{W}_2 \rangle$ is a *p-morphism* if f is a p-morphism from F_1 onto F_2, while $\mathbf{W}_1 \supseteq \{f^{-1}[X] \mid X \in \mathbf{W}_2\}$.

The relevant results about truth of modal formulas in general frames (corresponding to 2.11, 2.12, 2.15, 2.17 and 2.18) are obvious and will not be stated here.

General frames induce modal algebras, just like frames.

4.4 *Definition.* If $\langle \mathsf{F}, \mathbf{W} \rangle$ is a general frame, then the modal algebra $\langle \mathsf{F}, \mathbf{W} \rangle^+$ is $\langle \mathbf{W}, W, -, \cap, l \rangle$, where l is as defined above.

Conversely, general frames serve as set-theoretic representations for modal algebras through the well-known Stone representation:

4.5 *Definition.* If $\mathbf{A}$ is a modal algebra, then $SR(\mathbf{A})$ is the general frame $\langle \mathsf{F}, \mathbf{W} \rangle$ $(= \langle \langle W, R \rangle, \mathbf{W} \rangle)$, where

(i) W is the set of all ultrafilters on $\mathbf{A}$
(ii) RU_1U_2 if, for all $a \in A$, $l(a) \in U_1$ only if $a \in U_2$
(iii) $\mathbf{W}$ is the set of all subsets of W of the form $\{U \mid a \in U\}$ for some $a \in A$.

The usual isomorphism connects $\mathbf{A}$ with $SR(\mathbf{A})^+$. But, in order that a duality obtain between the categories of general frames and modal algebras, it is necessary also to have the following isomorphism: $\langle \mathsf{F}, \mathbf{W}\rangle \cong SR(\langle \mathsf{F}, \mathbf{W}\rangle^+)$. This is not true for all general frames, however; whence the following definition.

4.6 *Definition.* A general frame $\langle \mathsf{F}, \mathbf{W}\rangle$ is *descriptive* if it satisfies the following three conditions.

(i) For any $w, v \in W$, if $w \neq v$, then, for some $X \in \mathbf{W}$, $w \in X$ and $v \notin X$

(ii) For any $w, v \in W$, if not Rwv, then, for some $X \in \mathbf{W}$, $w \in \mathrm{l}(X)$ and $v \notin X$

(iii) Any set $S \subseteq \mathbf{W}$ with the finite intersection property has a non-empty intersection.

Conditions (i) and (ii) may be restated more elegantly, using the fact that $\langle \mathsf{F}, \mathbf{W}\rangle$ is a (general) model for L_2 as well:

(i) $\langle \mathsf{F}, \mathbf{W}\rangle \models \forall x \forall y(\forall P(Px \rightarrow Py) \rightarrow x = y)$
(identity of indiscernibles)

(ii) $\langle \mathsf{F}, \mathbf{W}\rangle \models \forall x \forall y(\forall P(\forall z(Rxz \rightarrow Pz) \rightarrow Py) \rightarrow Rxy)$.

The importance of descriptive general frames lies in the fact that they are the *fixed points* of the Stone representation:

4.7 *Lemma* A general frame $\langle \mathsf{F}, \mathbf{W}\rangle$ is descriptive iff $\langle \mathsf{F}, \mathbf{W}\rangle \cong SR(\langle \mathsf{F}, \mathbf{W}\rangle^+)$.

We now turn to the parallel between the main algebraic concepts and those concerning general frames, on which the above duality rests.

4.8 *Lemma.* If f is a homomorphism from the modal algebra $\mathbf{A}_1$ onto $\mathbf{A}_2$, then the function g from $SR(\mathbf{A}_2)$ into $SR(\mathbf{A}_1)$ defined by $g(U) = f^{-1}[U]$ is an isomorphism onto a generated subframe of $SR(\mathbf{A}_1)$.

If $\langle \mathsf{F}_1, \mathbf{W}_1\rangle$ is a generated subframe of $\langle \mathsf{F}_2, \mathbf{W}_2\rangle$, then the function g from $\langle \mathsf{F}_2, \mathbf{W}_2\rangle^+$ onto $\langle \mathsf{F}_1, \mathbf{W}_1\rangle^+$ defined by $g(X) = X \cap W_1$ is a homomorphism.

4.9 *Lemma.* If the modal algebra $\mathbf{A}_1$ is a subalgebra of $\mathbf{A}_2$, then the function g from $SR(\mathbf{A}_2)$ onto $SR(\mathbf{A}_1)$ defined by $g(U) = U \cap A_1$ is a p-morphism.

If f is a p-morphism from the general frame $\langle \mathbf{F}_1, \mathbf{W}_1\rangle$ onto $\langle \mathbf{F}_2, \mathbf{W}_2\rangle$, then the function g from $\langle \mathbf{F}_2, \mathbf{W}_2\rangle^+$ to $\langle \mathbf{F}_1, \mathbf{W}_1\rangle^+$ defined by $g(X) = f^{-1}[X]$ is an isomorphism onto a subalgebra of $\langle \mathbf{F}_1, \mathbf{W}_1\rangle^+$.

4.10 *Lemma.* The function g defined by $g(X) = \langle X \cap W_i\rangle_{i\in I}$ is an isomorphism between $(\Sigma\{\langle \mathbf{F}_i, \mathbf{W}_i\rangle \mid i \in I\})^+$ and $\Pi\{\langle \mathbf{F}_i, \mathbf{W}_i\rangle^+ \mid i \in I\}$.

$SR(\Pi\{\mathbf{A}_i \mid i\in I\})$ is not always isomorphic to $\Sigma\{SR(\mathbf{A}_i) \mid i \in I\}$, however. By 4.8, it does hold that, since $\mathbf{A}_i$ is a homomorphic image of $\Pi\{\mathbf{A}_i \mid i\in I\}$, $SR(\mathbf{A}_i)$ is isomorphic to a generated subframe of $SR(\Pi\{\mathbf{A}_i \mid i\in I\})$, and hence $\Sigma\{SR(\mathbf{A}_i) \mid i \in I\}$ is isomorphic to a generated subframe of $SR(\Pi\{\mathbf{A}_i \mid i\in I\})$ as well. Moreover, it does hold that $SR(\Pi\{\mathbf{A}_i \mid i\in I\})$ is isomorphic to $SR((\Sigma\{SR(\mathbf{A}_i) \mid i \in I\})^+)$.

These three results will enable us to use Birkhoff's theorem characterizing the *equational varieties* (i.e., the classes of algebras defined by some set of equations) as the classes of algebras closed under homomorphisms, subalgebras and direct products. (Cf. chapters 14 and 16.) It should be remarked, however, that the duality extends further than this, and that, e.g., Birkhoff's results about *subdirectly irreducible* algebras may be applied as well (cf. Blok [14]). Note the similarity between lemma 2.20 and Birkhoff's theorem saying that any algebra is isomorphic to a subalgebra of a product of subdirectly irreducible algebras. In fact, for any general frame $\langle \mathbf{F}, W\rangle$ such that $\mathbf{F}$ is generated, $\langle \mathbf{F}, W\rangle^+$ is a subdirectly irreducible algebra. The converse does not hold, however; but the arising complications are not relevant to the present work.

One important notion concerning general frames remains to be introduced, viz. that of "ultraproduct". It is clear from classical model theory how ultraproducts of *frames* (in the sense of chapter 2) or *modal algebras* are to be defined (cf. [17]). The case of general frames calls for some reflection, however.

4.11 *Definition.* Let $\{\langle \mathbf{F}_i, W_i\rangle \mid i\in I\}$ be a family of general frames, and U an ultrafilter on I. The *ultraproduct* $\Pi_U\{\langle \mathbf{F}_i, W_i\rangle \mid i\in I\}$ of this family is the general frame $\langle \mathbf{F}, W\rangle$ with

(i) $\mathbf{F}$ $(= \langle W, R\rangle)$ is the ordinary ultraproduct $\Pi_U\{\mathbf{F}_i \mid i\in I\}$, and

(ii) W consists of all subsets of W which are of the form $\Pi_U\{X_i \mid i\in I\} = \{f_U\in W \mid \{ i\in I \mid f(i)\in X_i\}\in U\}$; where, for each $i \in I$, $X_i\in W_i$.

The verification that $\Pi_U\{\langle \mathbf{F}_i, W_i\rangle \mid i \in I\}$ is a general frame is routine: the three closure conditions required by definition 4.2 hold in the ultraproduct because they hold in all factors. E.g., the clause for the operation l follows from the equivalence:

$$\exists\, g_U \in W(Rf_U g_U \ \&\ g_U \in \Pi_U\{X_i \mid i \in I\}) \text{ if and only if } f_U \in \Pi_U\{\{w \in W_i \mid \exists\, v \in X_i\ R_i wv\} \mid i \in I\};$$

which is proved just like the existential quantifier step in the proof of the fundamental Theorem of Łos.

Before returning to algebra, let us note that such an ultraproduct of general frames is still a (general) model for the second-order language L_2. Thus, the following generalization of Łos' Theorem becomes available; which will be used in chapters 13, 14 and 17.

4.12 *Theorem.* For any formula φ of the monadic second-order language based upon L_0 (i.e., L_2) with identity, having the free individual variables $x_1, \ldots, x_k$ and the free, unary predicate variables $P_1, \ldots, P_m$,

$$\Pi_U\{\langle \mathbf{F}_i, W_i\rangle \mid i \in I\} \models \varphi\, [f^1_U, \ldots, f^k_U;\ \Pi_U\{X^1_i \mid i \in I\}, \ldots, \Pi_U\{X^m_i \mid i \in I\}]$$
if and only if $\{i \in I \mid \langle \mathbf{F}_i, W_i\rangle \models \varphi[f^1(i), \ldots, f^k(i);\ X^1_i, \ldots, X^m_i]\} \in U$.

Proof. The proof of this assertion is by induction with respect to the complexity of L_2-formulas. It is completely analogous to the first-order case (cf. [17] or [30]). QED.

In the last analysis, this whole matter boils down to Łos' Theorem for many-sorted first-order logic (cf. chapter 17, or [20]).

Note the following important distinction. From a family of frames $\{\mathbf{F}_i \mid i \in I\}$ one obtains "ordinary" ultraproducts $\Pi_U\{\mathbf{F}_i \mid i \in I\}$, which may be regarded as general frames in the usual way; i.e. using the power set of the resulting domain. But, if the factors $\mathbf{F}_i$ themselves are regarded as general frames in this way, say as $\langle \mathbf{F}_i, P(W_i)\rangle$, then the ultraproducts $\Pi_U\{\langle \mathbf{F}_i, P(W_i)\rangle \mid i \in I\}$ as defined in 4.11 will normally be different from the former: they contain far fewer sets. (It is this poverty which enabled us to prove theorem 4.12.) E.g., if a modal formula φ is true in all $\mathbf{F}_i (i \in I)$, then it will be true in the latter ultraproduct — by the above equivalence — but it may fail to be true in the former ultraproduct. Preservation under ultraproducts in the former sense is even equivalent to first-order definability, as will be shown in chapter 8.

Finally, there is a useful algebraic duality result which is rather like lemma 4.10 (cf.[30], corollary 7.8):

4.13 *Lemma.* The function g defined by $g(\Pi_U\{X_i \mid i \in I\}) = (\langle X_i \rangle_{i \in I})_U$ is an isomorphism between $(\Pi_U\{\langle \mathbf{F}_i, W_i \rangle \mid i \in I\})^+$ and the ultraproduct $\Pi_U\{\langle \mathbf{F}_i, W_i \rangle^+ \mid i \in I\}$ of the modal algebras $\langle \mathbf{F}_i, W_i \rangle^+$ $(i \in I)$.

These algebraic techniques will become important in parts III and IV of this book.

CHAPTER V

AXIOMATIC THEORIES

In this chapter, some basic information will be given about the syntactic side of modal logic. Since it is customary to use $\top, \neg, \rightarrow$ and L as primitives in this area, that convention will be followed here as well.

Consider some set of axiom schemas complete for propositional logic when the only rule of inference is *Modus Ponens* (i.e., to infer ψ from φ and $\varphi \rightarrow \psi$). The *minimal modal logic* K arises from the addition of the following principles (for all modal formulas φ and ψ).

Definition: $M\varphi = \neg L \neg \varphi$
Axiom Schema: $L(\varphi \rightarrow \psi) \rightarrow (L\varphi \rightarrow L\psi)$
Rule of inference: to infer $L\varphi$ from φ (*Necessitation*).

Derivability within K may now be defined in the obvious manner. (Notation: $\Sigma \vdash_K \varphi$ for "φ is derivable from Σ within K".)

A well-known variant of K is obtained by taking single axioms instead of axiom schemas and adding a *Rule of Substitution* (i.e., to infer any substitution instance of a formula already obtained). The latter system will be called K_s. By a result of von Neumann's, K and K_s have exactly the same theorems. Again one writes $\Sigma \vdash_{K_s} \varphi$ for "φ is derivable from Σ within K_s".

The modal logics to be studied here all consist of K with additional axioms. Such axiomatic theories are called *normal* modal logics in the literature, to distinguish them from weaker (or incomparable) systems that are studied as well (cf. Segerberg [67]). A few normal modal logics will be mentioned here, to give an impression of the kind of modal axioms found in the literature.

Logic	*Characteristic axioms*
D	$M \top$
T	$L\varphi \to \varphi$
$S4$	$L\varphi \to \varphi$, $L\varphi \to LL\varphi$
$S4.1$	$L\varphi \to \varphi$, $L\varphi \to LL\varphi$, $LM\varphi \to ML\varphi$
$S4.2$	$L\varphi \to \varphi$, $L\varphi \to LL\varphi$, $ML\varphi \to LM\varphi$
$S4.3$	$L\varphi \to \varphi$, $L\varphi \to LL\varphi$, $L(L\varphi \to L\psi) \vee L(L\psi \to L\varphi)$
$S5$	$L\varphi \to \varphi$, $L\varphi \to LL\varphi$, $ML\varphi \to \varphi$
B	$L\varphi \to \varphi$, $ML\varphi \to \varphi$
Dum	$L(L(\varphi \to L\varphi) \to \varphi) \to \varphi$
LF	$L(L\varphi \to \varphi) \to L\varphi$
Id	$\varphi \leftrightarrow L\varphi$
Un	$L\bot$

These are *modal* axioms. In other branches of intensional logic (such as *tense* logic or *deontic* logic) different principles will be needed. E.g., $L\varphi \to \varphi$ — a modal triviality — is very implausible in deontic logic (where "L" is interpreted as "*obligatory*") and downright false in tense logic (where "L" is interpreted as "(from now on) *always in the future*").

In order to show what modal deduction looks like, two recent, ingenious proofs will be given. Recall that LF implied transitivity (chapter 3, example (9)). D.H.J. de Jongh showed (and G. Sambin discovered this independently, cf. [54]) that $Lp \to LLp$ is derivable in LF:

(1) Substitute $p \wedge Lp$ for p in LF. This yields the axiom $L(L(p \wedge Lp) \to (p \wedge Lp)) \to L(p \wedge Lp)$.

(2) $Lp \to L(L(p \wedge Lp) \to (p \wedge Lp))$ is provable in LF; in fact, it is derivable in K already:
$L(p \wedge Lp) \to Lp$, $p \to (L(p \wedge Lp) \to (p \wedge Lp))$, $L(p \to (L(p \wedge Lp) \to (p \wedge Lp)))$, $Lp \to L(L(p \wedge Lp) \to (p \wedge Lp))$.

(3) From (1) and (2) it follows that $Lp \to L(p \wedge Lp)$ is provable in LF. This formula implies $Lp \to LLp$ in K and hence in LF.

Similarly, *Dum* implies transitivity (and reflexivity). It is trivial to show that $Lp \to p$ is derivable in *Dum*, but it took a long time before W. J. Blok (and independently K. Pledger) found a proof that $Lp \to LLp$ is derivable in the logic *Dum* (cf.[13]):

(1) Substitute $\sigma = p \wedge (Lp \to LLp)$ for p in *Dum*. This yields the

axiom $L(L(\sigma \to L\sigma) \to \sigma) \to \sigma$.

(2) $p \to (L(\sigma \to L\sigma) \to \sigma)$ is provable in *Dum*; in fact, it is derivable in K already:
$p \to (L(\sigma \to L\sigma) \to p)$ is trivially provable, and the following chain of inference proves $L(\sigma \to L\sigma) \to (Lp \to LLp)$: $(Lp \wedge \neg LLp) \to (Lp \wedge M\neg Lp) \to M(\neg Lp \wedge p) \to M(\neg Lp \wedge p \wedge (Lp \to LLp)) \to$ (using the antecedent formula $L(\sigma \to L\sigma)) \to M(\neg Lp \wedge L(p \wedge (Lp \to LLp))) \to M(\neg Lp \wedge Lp) \to M\bot \to \bot$. In other words, $(Lp \wedge \neg LLp) \to (L(\sigma \to L\sigma) \to \bot)$, or $L(\sigma \to L\sigma) \to \neg(Lp \wedge \neg LLp)$ (i.e., $Lp \to LLp$).

(3) From (2) it follows that $L(p \to L(\sigma \to L\sigma) \to \sigma))$ and, therefore, $Lp \to L(L(\sigma \to L\sigma) \to \sigma)$ are derivable in *Dum*. Combining this with (1) yields the provability of $Lp \to (p \wedge (Lp \to LLp))$. This formula implies $Lp \to (Lp \to LLp)$ and, therefore, $Lp \to LLp$ in K, and hence also in *Dum*.

The complexity of these modal deductions contrasts very unfavourably with the straightforward semantic arguments of chapter 3.

A multitude of syntactic results about modal logics has been found in the past decades, such as the normal form theorem for *S5* (cf. Hughes & Cresswell [36]), or the theorem showing that for any *closed* modal formula φ (i.e., φ contains no proposition letters) $\vdash_D \varphi$ or $\vdash_D \neg\varphi$ (cf.Segerberg [70]). An important result in this area is the syntactic counterpart to lemma 2.15: any consistent normal modal logic is either contained in *Id* or in *Un* (cf. Makinson [57]). The limitations on these results are shown by the fact that one soon needs to establish *un*provability as well. E.g. when proving that the number of modalities in *S4* is exactly 14 (cf. Hughes & Cresswell [36]), one not only has to prove certain equivalences (like $LMLMp \leftrightarrow LMp$) in *S4*, but also that other equivalences (like $LMLp \leftrightarrow MLp$) are not provable in this theory. For this purpose, one usually needs the method of *counterexamples*, i.e., a connection with semantics. This connection will be treated in the next chapter.

To facilitate the study of axiomatic theories like the above, the following formal definition is given.

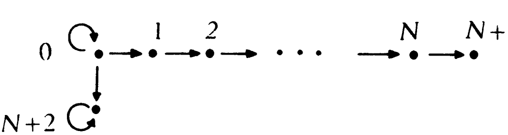
5.1 *Definition.* A *modal logic* is a set of modal formulas which contains all instances of axioms in K and is closed under Modus Ponens, Necessitation and the Rule of Substitution. For any set Σ of modal formulas, the modal logic *axiomatized by* Σ ($ML(\Sigma)$) is the smallest modal logic containing Σ.

The above examples were all axiomatized by means of *finite* sets of modal formulas. Unfortunately, the finitely axiomatizable modal logics do not form a nice subclass of the class of all modal logics. It is easily seen that, if φ_1 axiomatizes Σ_1 and φ_2 Σ_2, then the modal logic axiomatized by $\Sigma_1 \cup \Sigma_2$ is also axiomatized by $\varphi_1 \wedge \varphi_2$. But, the modal logic $\Sigma_1 \cap \Sigma_2$ need not be finitely axiomatizable at all! A counterexample is provided by $M\top$ and $Mp \rightarrow Lp$: $ML(\{M\top\}) \cap ML(\{Mp \rightarrow Lp\})$ is not finitely axiomatizable.

To see this, note first that, for any two modal formulas $\varphi_1, \varphi_2, ML(\{\varphi_1\}) \cap ML(\{\varphi_2\})$ is axiomatized by the infinite set $\{L^i\varphi_1 \vee L^j\varphi_2 \mid i, j \geqslant 0\}$, where $L^k\varphi = L$ ...(k times)... $L\varphi$. If this modal logic is axiomatized by some modal formula ψ, then it follows that, for some N, $\psi \in \mathrm{ML}(\{L^i\varphi_1 \vee L^j\varphi_2 \mid 0 \leqslant i, j \leqslant N\})$; whence $\{L^i\varphi_1 \vee L^j\varphi_2 \mid 0 \leqslant i, j \leqslant N\}$ axiomatizes the whole set. So, in particular, $L^{N+1}\varphi_1 \vee L^{N+1}\varphi_2 \in ML(\{L^i\varphi_1 \vee L^j\varphi_2 \mid 0 \leqslant i, j \leqslant N\})$. When this insight is applied to the above example it clearly suffices to show that, for no N, $L^{N+1}M\top \vee L^{N+1}(Mp \rightarrow Lp) \in ML(\{L^iM\top \vee L^j(Mp \rightarrow Lp) \mid 0 \leqslant i, j \leqslant N\})$. A counterexample will be constructed here, whose justification follows from the completeness theory of the next chapter.

Let $\mathbf{F} = \langle W, R \rangle$, where

$W = \{0, ..., N+2\}$,
$R = \{\langle i, i+1 \rangle \mid 0 \leqslant i \leqslant N\} \cup \{\langle 0, 0 \rangle, \langle 0, N+2 \rangle, \langle N+2, N+2 \rangle\}$.



$L^{N+1}M\top \vee L^{N+1}(Mp \rightarrow Lp)$ does not hold in $\mathbf{F}$. For, if $V(p)$ is taken to be $\{1\}$, then $\langle \mathbf{F}, V \rangle \models Mp \wedge \neg Lp[0]$; whence $L^{N+1}(Mp \rightarrow Lp)$ fails to hold at 0. Moreover, since $M\top$ fails at $N+1$, $L^{N+1}M\top$ does not hold at 0, even independently of the choice of a valuation. But $L^iM\top \vee L^j(Mp \rightarrow Lp)$ holds in $\mathbf{F}$ for any i, j with $0 \leqslant i, j \leqslant N$. To see this, check that

(i) $\mathbf{F} \models L^iM\top[N+2]$ for all $i \geqslant 0$,
(ii) $\mathbf{F} \models L^i(Mp \rightarrow Lp)[k]$ for all $i \leqslant 0$ and each $k \in \{1, ..., N+1\}$, and
(iii) $\mathbf{F} \models L^iM\top[0]$ for each $i \leqslant N(!)$.

What has been proven just now may be described algebraically as follows. The finitely axiomatizable modal logics do not form a *sublattice* of the lattice of all modal logics under the usual operations, viz. intersection and deductive closure of unions. (But note that, if attention is restricted to modal logics *containing* $S4$, then this result does hold. For, then $L\varphi_1 \quad L\varphi_2$ axiomatizes $ML(\{\varphi_1\}) \cap ML(\{\varphi_2\})$.) This negative result makes it preferable to study modal logics with *arbitrary* sets of axioms; as in universal algebra, where a restriction to algebraic theories axiomatized by only finitely many equations would endanger the elegance of the Birkhoff theory.

CHAPTER VI

COMPLETENESS

The notion $\vdash_K$ of derivability within the smallest modal logic K turns out to be the syntactic counterpart of the notion $\models_m$ of semantic consequence introduced in chapter 2:

6.1 *Theorem.* For all sets Σ of modal formulas and all modal formulas φ, $\Sigma \vdash_K \varphi$ iff $\Sigma \models_m \varphi$.

Note that the interesting fact about 6.1 is *not* that $\models_m$ has a syntactic characterization: this much was clear already from the translation of modal formulas into L_1 given in chapter 3. But it is useful to know that a simple system like K suffices.

Proof of 6.1 Theorem 6.1 is proven by means of a *Henkin construction* invented by D. C. Makinson (cf. [56]) and E. J. Lemmon & D. Scott (cf. [48]). Although it is well-known by now, some striking features will be repeated here. A routine induction with respect to the length of the derivation of φ from Σ establishes the direction from left to right. For the other direction, one considers the *Henkin Model* $HM(\Sigma)$ of Σ, consisting of the *Henkin Frame* $HF(\Sigma)$ of Σ together with a valuation V_Σ; which are defined as follows. $HF(\Sigma) = \langle W_\Sigma, R_\Sigma \rangle$, where W_Σ consists of all sets of modal formulas which are *maximally* Σ-*consistent* (Δ is Σ-*consistent* if, for no $\delta_1, ..., \delta_n \in \Delta$, $\Sigma \vdash_K \neg(\delta_1 \wedge ... \wedge \delta_n)$), $R_\Sigma \Delta_1 \Delta_2$ holds if, for all modal formulas φ, $L\varphi \in \Delta_1$ only if $\varphi \in \Delta_2$, and $V_\Sigma(p)$ is defined as $\{\Delta \in W_\Sigma \mid p \in \Delta\}$.

Note the similarity between the definition of R_Σ and that of R_F in

2.24. A proof essentially like that of 2.25 establishes the following equivalence for all modal formulas φ and all $\Delta \in W_\Sigma$:

$$HM(\Sigma) \models \varphi[\Delta] \text{ iff } \varphi \in \Delta.$$

Now, assume that $\Sigma \not\vdash_K \varphi$, i.e., that $\{\neg\varphi\}$ is Σ-consistent. By Lindenbaum's extension theorem, $\{\neg\varphi\}$ is contained in some $\Delta \in W_\Sigma$. Then, by the equivalence given above, $HM(\Sigma) \models \neg\varphi[\Delta]$; i.e., φ does not hold in $HM(\Sigma)$. But, clearly, $HM(\Sigma) \models \Sigma$, since every maximally Σ-consistent set contains Σ. So, $HM(\Sigma)$ provides a counterexample to $\Sigma \models_m \varphi$. QED.

Remark. The analogy between the proof of 6.1 and that of 2.25 —which was inspired by the algebraic Stone representation — is not surprising. Henkin-type proofs are nothing more — speaking algebraically, that is —than a combination of an algebraic completeness proof and a set-theoretic representation of the verifying Lindenbaum algebra.

Henkin models of the form $HM(\Sigma)$ were characterized by K. Fine as the only models satisfying certain properties called "tightness", "1-saturatedness" and "2-saturatedness" (cf. [24]). These properties express, for models, the requirements (i), (ii) and (iii) on descriptive general frames given in 4.6. The latter formulation can be used in this case as well, because Henkin models $HM(\Sigma)$ may be regarded as general frames $HGF(\Sigma)$ through the following reformulation (already mentioned in chapter 4): $HGF(\Sigma) = \langle HF(\Sigma), \mathcal{W}_\Sigma\rangle$, where $\mathcal{W}_\Sigma$ is defined as the set of all subsets of W_Σ which are of the form $\{\Delta \in W_\Sigma \mid \varphi \in \Delta\}$ for some modal formula φ. Note that, if Σ is closed under the formation of substitution instances, then $HGF(\Sigma) \models \Sigma$. This remark enables us to find a semantic equivalent to the notion of derivability within K_s ($\vdash_{K_s}$) introduced in chapter 5:

6.2 *Definition.* For a set Σ of modal formulas and a modal formula φ, $\Sigma \models_{gf} \varphi$ if, for all general frames $\langle \mathbf{F}, \mathcal{W}\rangle$, $\langle \mathbf{F}, \mathcal{W}\rangle \models \Sigma$ only if $\langle \mathbf{F}, \mathcal{W}\rangle \models \varphi$.

Note that $\Sigma \models_m \varphi$ only if $\Sigma \models_{gf} \varphi$ only if $\Sigma \models_f \varphi$. None of these implications may be reversed, however. For the first one, this is trivial: $\{p\} \models_{gf} \perp$, but not $\{p\} \models_m \perp$. A counterexample to the second implication will be given below. An easy reformulation of the above proof yields the following completeness theorem.

6.3 *Theorem.* For all sets Σ of modal formulas and all modal formulas φ, $\Sigma \vdash_{Ks} \varphi$ (i.e., $\varphi \in ML(\Sigma)$) iff $\Sigma \models_{gf} \varphi$.

To repeat, general frames of the form $HGF(\Sigma)$ are descriptive. (Moreover, conversely, any descriptive general frame may be viewed as (an isomorphic copy of) a general frame $HGF(\Sigma)$, for some suitably chosen Σ in some suitably constructed modal propositional language.)

The better-known modal completeness theorems are of yet a different form, however. E.g., a modal formula is in $S5$ if and only if it holds in all frames whose alternative relation is an equivalence relation. If this is restated in the present terms, using the fact (cf. chapter 3) that the characteristic axioms $Lp \rightarrow p$, $Lp \rightarrow LLp$ and $MLp \rightarrow p$ of $S5$ define reflexivity, transitivity and symmetry — in that order —, then it assumes the following form. For the set Σ axiomatizing $S5$ and for any modal formula φ,

$$\varphi \in ML(\Sigma) \text{ iff } \Sigma \models_f \varphi;$$

where $\models_f$ is the relation of semantic consequence on the class of all frames introduced in chapter 2. In other words, modal completeness theorems are about sets Σ for which $\models_f$ and $\vdash_{Ks}$ coincide.

6.4 *Definition.* A set Σ of modal formulas is *complete* ($\Sigma \in C$) if, for all modal formulas φ, $\varphi \in ML(\Sigma)$ iff $\Sigma \models_f \varphi$.

Another way to phrase this is as follows.

6.5 *Lemma.* Σ is complete iff there exists a frame $\mathbf{F}$ such that $Th_{mod}(\mathbf{F}) = ML(\Sigma)$.

Proof. The direction from right to left is obvious. For the converse direction, take an $\mathbf{F}$ for every $\varphi \notin ML(\Sigma)$ such that $\mathbf{F} \models \Sigma$ and $\mathbf{F} \not\models \varphi$, and then consider the disjoint union of these $\mathbf{F}$'s. QED.

Results such as the completeness of $S5$ are proven quite easily, by inspection of $HM(S5)$. It suffices to show that $S5$ does not only hold on $HGF(S5)$, but also on the underlying frame $HF(S5)$; because that frame falsifies all non-theorems of $S5$. Now, an easy argument shows that the characteristic axioms of *S5* induce the relational properties they define (thanks to the definition of R_{S5}). In many cases, completeness proofs require additional techniques, however, such as the use of *filtrations* (cf. Segerberg [67]). Although such proofs are often quite ingenious, a lack of general methods and results was conspicuous before 1974. (One of the notable exceptions was *Bull's Theorem* stating that every extension of *S4.3* is complete. Cf. also K. Fine's two papers [23] and [25].) Therefore, people began to lose the hope that $\vdash_{K\Sigma}$ would axiomatize $\models_f$ (i.e., that Kripke semantics as originally intended would suffice for modelling each modal logic).

In 1974 K. Fine (cf. [22]) and S. K. Thomason (cf. [75]) gave examples of *incomplete* modal logics. I.e., they gave Σ and φ such that $\Sigma \models_f \varphi$, but not $\varphi \in ML(\Sigma)$. (The other way around is impossible: if $\varphi \in ML(\Sigma)$, then $\Sigma \models_f \varphi$.) Here we mention a simple example taken from [13]: Let $\Sigma = \{L(Lp \rightarrow Lq) \vee L(Lq \rightarrow Lp),\ Lp \rightarrow p,\ LMp \rightarrow MLp,\ (Mp \wedge L(p \rightarrow Lp)) \rightarrow p\}$. $\Sigma \models_f p \rightarrow Lp$ (i.e., Σ holds on exactly the same frames as the "classical" logic *Id* axiomatized by $p \leftrightarrow Lp$), but $p \rightarrow Lp \in ML(\Sigma)$ (i.e., $ML(\Sigma)$ is a proper subset of *Id*).

That $p \rightarrow Lp \notin ML(\Sigma)$ is shown by exhibiting a general frame where Σ holds, but $p \rightarrow Lp$ fails.

Another incompleteness example is simple enough to be explained here (cf. [12]). Let $\Sigma = \{LM \quad \rightarrow L(L(Lp \rightarrow p) \rightarrow p)\}$ and $\varphi = LM\top \rightarrow L\bot$.

6.6 *Theorem.* $LM\top \rightarrow L(L(Lp \rightarrow p) \rightarrow p) \models_f LM\top \rightarrow L\bot$; but $LM\top \rightarrow L\bot \in ML(\{LM\top \rightarrow L(L(Lp \rightarrow p) \rightarrow p)\})$.

Proof. The first assertion is a short exercise in L_2-deduction. Observe that

$$\forall P\ ST(L(L(Lp \rightarrow p) \rightarrow p)) =$$
$$\forall P \forall y(Rxy \rightarrow (\forall z(Ryz \rightarrow (\forall u(Rzu \rightarrow Pu) \rightarrow Pz)) \rightarrow Py))$$

implies its L_0-substitution instance (replacing subformulas "$P\star$" by "$\star \neq y$"):

$$\forall y(Rxy \rightarrow (\forall z(Ryz \rightarrow (\forall u(Rzu \rightarrow u \neq y) \rightarrow z \neq y) \rightarrow y \neq y)).$$

The latter L_0-formula is equivalent to

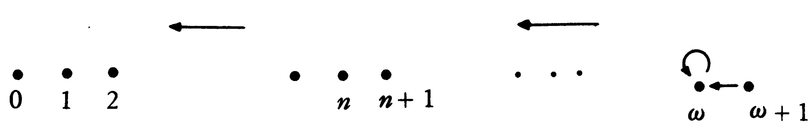
$\forall y(Rxy \rightarrow \neg\forall z(Ryz \rightarrow (\forall u(Rzu \rightarrow u \neq y) \rightarrow z \neq y)))$, and hence to

$\forall y(Rxy \to \exists z(Ryz \wedge \forall u(Rzu \to u \neq y) \wedge z = y))$, i.e., to
$\forall y(Rxy \to (Ryy \wedge \forall u(Ryu \to u \neq y))$, i.e., to
$\forall y(Rxy \to (Ryy \wedge \neg Ryy))$, i.e., to
$\neg \exists y Rxy = ST(L\bot)$.

This observation shows that $ST(LM\top \to L\bot)$ may even be derived from $\forall P\ ST(LM\top \to L(L(Lp \to p) \to p))$ in a very weak second-order logic, notably one without the axiom of choice. (The latter axiom was used in all former examples of incompleteness.) This makes it all the more interesting that K_s fails to provide for a proof of this implication. Here is the counter-example.

Let $\langle \mathbf{F}, \mathbf{W} \rangle$ be the general frame given by:

$W = I\!N \cup \{\omega, \omega + 1\}$,
$R = \{\langle i, j \rangle \mid i, j \in I\!N \ \&\ i > j\} \cup \{\langle \omega, i \rangle \mid i \in I\!N\} \cup \{\langle \omega, \omega \rangle, \langle \omega + 1, \omega \rangle\}$,
and
$\mathbf{W} = \{X \subseteq W \mid (X \text{ is cofinite } \&\ \omega \in X) \text{ or } (X \text{ is finite } \&\ \omega \notin X)\}$.



It is easily verified that $\mathbf{W}$ satisfies the three closure conditions for general frames. Clearly, $LM\top \to L\bot$ does not belong to $Th_{mod}(\langle \mathbf{F}, \mathbf{W} \rangle)$: consider the world $\omega + 1$. But, $LM\top \to L(L(Lp \to p) \to p)$ is true in this general frame. For, if V is any valuation assigning sets in $\mathbf{W}$ to the proposition letters such that $\langle \mathbf{F}, V \rangle \models LM\top[w]$, then w must be $\omega + 1$; since all other worlds have the end-point 0 as an R-successor. Thus, it suffices to show that $\langle \mathbf{F}, V \rangle \models L(Lp \to p) \to p[\omega]$, ω being the only R-successor of $\omega + 1$. Now, if $\langle \mathbf{F}, V \rangle \models L(Lp \to p)[\omega]$, then $\langle \mathbf{F}, V \rangle \models Lp \to p[0]$, and hence $\langle \mathbf{F}, V \rangle \models p[0]$ (since Lp is true at any end-point). But, then, p is true at all R-successors of 1; and hence $\langle \mathbf{F}, V \rangle \models Lp[1]$. Moreover, the truth of $L(Lp \to p)$ at ω implies that $\langle \mathbf{F}, V \rangle \models Lp \to p[1]$, whence $\langle \mathbf{F}, V \rangle \models p[1]$. Continuing this argument with respect to 2, etc., one shows successively that each natural number belongs to $V(p)$. It follows that $V(p)$ is cofinite. Now, since $V(p) \in \mathbf{W}$, this means that $\omega \in V(p)$. QED.

Of course, by S. K. Thomason's result mentioned in chapter 1, the general background of these incompleteness phenomena is clear: $\models_f$ is essentially an (unaxiomatizable) second-order notion of consequence. The introduction of general frames relaxed this notion to an axiomatizable one in the same way as Henkin's introduction of general models yielded an axiomatization of (weakened) second-order consequence (cf. Henkin [33]).

A closer look at Segerberg's [67] shows that modal completeness theorems are never formulated in the abstract way of definition 6.4:

6.7 *Definition.* A set Σ of modal formulas is *complete with respect to* the class $\mathbf{K}$ of frames if, for all modal formulas φ, $\varphi \in ML(\Sigma)$ iff $\varphi \in Th_{mod}(\mathbf{K})$ (i.e., $\forall \mathbf{F} \in \mathbf{K}$: $\mathbf{F} \models \varphi$).

E.g., *S5* is complete with respect to the class of frames whose relation is an equivalence relation. But, it is also complete with respect to the class of frames whose relation is the *universal* relation on their domains (i.e., $\forall x \forall y Rxy$ holds). This strengthening follows from the generation theorem (2.11): if φ is falsified at w in a frame $\mathbf{F} = \langle W, R \rangle$, where R is an equivalence relation, then it is also falsified in $TC(\mathbf{F}, w)$, whose domain is the R-equivalence class of w; and, on that equivalence class, R is universal. Note, however, that, by the same generation theorem, $\forall x \forall y Rxy$ was shown to be undefinable by *S5* (or any other set of modal formulas).

If Σ is complete with respect to any class of frames, then it is complete with respect to $FR(\Sigma)$, i.e., it is complete in the sense of 6.4. But the formulation given in 6.7 suggests another point of view as well: a class $\mathbf{K}$ is given and one asks for a recursive axiomatization of the modal logic $Th_{mod}(\mathbf{K})$. (Any such axiomatization will be complete with respect to $\mathbf{K}$.) Note that, if $\mathbf{K}$ is *elementary* (say the L_0-sentence α defines it), then $Th_{mod}(\mathbf{K}) = \{\varphi \mid \alpha \models \forall x ST(\varphi)\}$ is recursively enumerable and, therefore — by Craig's Theorem — a recursive modal axiomatization exists for this set.

Finally, a convenient syntactic point of view concerning completeness should be mentioned; which is due to R. I. Goldblatt. For any set of modal formulas, a kind of "complete closure" may be defined as follows.

6.8 *Definition.* Let Σ be a set of modal formulas. $C(\Sigma)$ is the set $Th_{mod}(FR(\Sigma))$.

It is trivial that $\Sigma \subseteq C(\Sigma)$; but the converse is more interesting. In fact, some calculation shows that the fixed points of the operation C are precisely the complete modal logics! Moreover, $C(\Sigma)$ is the smallest complete modal logic to contain Σ. As has been shown by W. J. Blok in recent work, there exist even uncountable sets of distinct modal logics all sharing the same complete closure.

Some special classes of complete sets are the following.

6.9 *Definition.* Σ is *generally complete* ($\Sigma \in GC$) if, for any modal logic Δ containing Σ and any modal formula φ, $\varphi \in \Delta$ iff, for all general frames $\langle \mathbf{F}, \mathbf{W} \rangle$ such that $\mathbf{F} \models \Sigma$ and $\langle \mathbf{F}, \mathbf{W} \rangle \models \Delta$, $\langle \mathbf{F}, \mathbf{W} \rangle \models \varphi$.

Clearly, any generally complete Σ is complete. An important kind of generally complete sets are the *canonical* ones.

6.10 *Definition* (cf. Fine [24]). Σ is *canonical* if, for any modal propositional language — no matter what the set of proposition letters is taken to be (countable or uncountable) — Σ holds in the Henkin frame $HF(\Sigma)$.

This definition of canonicity is not very elegant, being language-dependent. Therefore, it is replaced by the following one, involving a preservation condition.

6.11 *Definition.* Σ is *canonical* ($\Sigma \in CAN$) if, for any descriptive general frame $\langle \mathbf{F}, \mathbf{W} \rangle$ such that $\langle \mathbf{F}, \mathbf{W} \rangle \models \Sigma$, $\mathbf{F} \models \Sigma$.

Clearly, canonicity in the second sense implies canonicity in the first sense (because Henkin general frames are descriptive). The converse only holds if definition 6.10 is somewhat strengthened (cf. [3]).

Canonicity in the sense of 6.11 is called "d-persistence" in Goldblatt [30]. Canonical sets will reappear in chapters 13 and 16.

A final interesting class of complete sets is the following.

6.12 *Definition.* Σ is *first-order complete* ($\Sigma \in C1$) if a set Δ of L_0-sentences exists such that Σ is complete with respect to $FR(\Delta)$.

All definitions given above may be specialized to the case of single modal formulas. E.g., if 6.12 is reformulated to "$\varphi \in C1$ if, for some L_0-sentence α, $\psi \in ML(\{\varphi\})$ iff $\alpha \models \forall x ST(\psi)$, for all modal formulas ψ", then $\{\varphi \mid \varphi \in C1\}$ is an arithmetically definable class of modal formulas, which is — presumably (cf. chapter 7) — much less complex than $\bar{M}1$.

For more examples of subclasses of C, cf. chapter 13.

To conclude this chapter, some connections between the concepts introduced here (and those introduced in chapter 3) will be sketched. As we are concerned with sets of formulas, $\bar{M}1$ will be taken to be $\{\Sigma \mid$ for some set Δ of L_0-sentences, $FR(\Sigma) = FR(\Delta)\}$.

First, there are some obvious inclusions.

(1) $CAN \subseteq GC \subseteq C$
(2) $\bar{M}1 \cap C \subseteq C1$.

Fine [24] contains the formula $MLp \rightarrow (ML(p \wedge q) \vee ML(p \wedge \neg q))$ (or, equivalently, $ML(p \vee q) \rightarrow (MLp \vee MLq)$), which is in CAN (and in $C1$ as well), but which is not in $\bar{M}1$. In other words,

(3) $CAN \not\subseteq \bar{M}1$
(4) $C1 \not\subseteq \bar{M}1$.

Moreover, the following negative result holds.

(5) $C \not\subseteq C1$.

Proof. A counterexample is provided by LF $(L(Lp \rightarrow p) \rightarrow Lp)$, which is complete (cf. Segerberg [67]), but which is not complete with respect to any class $FR(\Delta)$, where Δ is some set of L_0-sentences. For, suppose it were. Since, for no n, $L^n\bot \in ML(\{LF\})$ (to see this, consider any finite linear order of length $> n$: LF holds in it, $L^n\bot$ does not), the following set of formulas would be finitely satisfiable:

$$\Delta \cup \{Rxx_1, Rx_1x_2, \ldots, Rx_{n-1}x_n, \ldots) \mid n \geqslant 1\}$$
$$\cup \{\forall y(x = y \vee Rxy), \forall x \forall y(x = y \vee Rxy \vee Ryx)\}.$$

By compactness, the whole set would be satisfiable, and so a frame $\mathbf{F}$ and $w \in W$ would exist such that $\mathbf{F} \models \Delta$ and some infinite ascending R-sequence starts from w. But, LF does not hold in such a frame: a contradiction, for $FR(\Delta) \subseteq FR(\Sigma)$. **QED.**

LF also provides an example of a formula which is in C, but not in

CAN. For, LF is not preserved under ultrafilter extensions (cf. chapter 2), while all canonical modal formulas are.

This last fact is proven as follows. Let $\varphi \in CAN$ and let $\mathbf{F} \models \varphi$. $SR(\mathbf{F}^+) = \langle ue(\mathbf{F}), W \rangle$ is a descriptive general frame in which φ holds (cf. chapter 4). But then, since $\varphi \in$ CAN, $ue(\mathbf{F}) \models \varphi$. We have shown:

(6) $C \not\subseteq CAN$.

Next come some less obvious inclusions than (1) and (2).

(7) $C1 \subseteq GC$.

Proof. Cf. M. Mortimer [61]. (A slight modification of Mortimer's proof will do.)

(8) $\bar{M}1 \cap C \subseteq CAN$.

Proof. Cf. K.Fine [24]. This inclusion will be shown to hold in chapter 16.

Especially interesting is the relation between CAN and $C1$. It follows from Fine's proof in [24] that, if Σ is complete with respect to $FR(\Delta)$ for some set Δ of L_0-sentences which is preserved under disjoint unions, then Σ is canonical. It was shown in [11] that

(9) $C1 \subseteq CAN$,

using a modification of the proof of theorem 8.9.

As for the converse, we hazard a *conjecture*:

(10) $CAN \not\subseteq C1$.

Finally, the connection between $\bar{M}1$ and C is far from clear (cf. [11]):

(11) $\bar{M}1 \not\subseteq C$

(12) $C \not\subseteq \bar{M}1$

(13) not all modal formulas are in $\bar{M}1 \cup C$.

Proof of (11). Cf. the incomplete modal logic given above, which defined $FR(p \leftrightarrow Lp)$, i.e., $FR(\forall x \forall y (Rxy \leftrightarrow x = y))$.

Proof of (12). LF is a counterexample: it is in C, but — as was shown in chapter 3 — it is not in $\bar{M}1$, because the relational property "transitivity of R + well-foundedness of the converse of R" is not first-order definable.

Proof of (13). An example of this is found in [13]:

$$\{L(L(p \rightarrow Lp) \rightarrow LLLp) \rightarrow p,\ Lp \rightarrow p\}$$

is incomplete and not first-order definable. (S. K. Thomason's formula δ of chapter 1 is another, but far more complicated, example.)

There is as yet no satisfactory characterization of the class $\bar{M}1 \cap C$. (But cf. chapter 13 for a line of attack.)

PART II

First-Order Definability of Modal Formulas

CHAPTER VII

LOCAL AND GLOBAL FIRST-ORDER DEFINABILITY

To begin with, let us recall some definitions and results from chapter 3. In 3.10, $E(\varphi, \alpha)$ was defined to hold between a modal formula φ and an L_0-formula α with one free variable if $\forall \mathbf{F}$ $(= \langle W, R \rangle)$ $\forall w \in W$ $(\mathbf{F} \models \varphi[w] \Leftrightarrow \mathbf{F} \models \alpha[w])$. This is the relation of *local equivalence*. There is also a relation $\bar{E}$ of *global equivalence* between φ and an L_0-sentence α, which holds if $\forall \mathbf{F}(\mathbf{F} \models \varphi \Leftrightarrow \mathbf{F} \models \alpha)$. It was noted that if $E(\varphi, \alpha)$, where α has the free variable x, then $\bar{E}(\varphi, \forall x \alpha)$. Corresponding to these relations, there were two notions of *first-order definability* for modal formulas; one local ($M1 = \{\varphi \mid$ for some α, $E(\varphi, \alpha)\}$) and one global ($\bar{M}1 = \{\varphi \mid$ for some α, $\bar{E}(\varphi, \alpha)\}$). It follows from the above connection between E and $\bar{E}$ that $M1 \subseteq \bar{M}1$. It will be shown below that $\bar{M}1 \not\subseteq M1$, however.

Depending upon one's semantical preferences, the local or global notions may seem more attractive. E.g., for people working with distinguished "actual worlds" in frames (as Kripke himself did originally), the local versions will be preferable. From a technical point of view, both have their uses. The global notions are more tractable and elegant when model theory is used, the local ones lend themselves more easily to syntactic investigations (cf. chapter 9).

When sets of modal formulas are considered instead of single ones, the situation changes. The above definitions extend easily to sets Σ, Δ (instead of φ, α). For sets which are closed under prefixing of L (like the modal logics of chapter 5), lemma 3.11 stated that $\Sigma \in M1$ if and only if $\Sigma \in \bar{M}1$. (It will be proven in chapter 11 that, on the class of *transitive* frames, $\varphi \in \bar{M}1$ if and only if $L\varphi \in M1$. This does not hold in general, however. By lemma 8.8, if $L\varphi \in M1$, then $\varphi \in M1$. Therefore, one would have that, for all $\varphi \in \bar{M}1$, ($L\varphi$ and so) $\varphi \in M1$; i.e., $\bar{M}1 \subseteq M1$: which is false). One might also consider "mixed" cases, such as φ, Δ; for which we had that, if φ is locally (globally) equivalent to some set Δ of L_0-formulas, then it is locally (globally) equivalent to a single such formula. The other mixed case, that of an L_0-formula α being equivalent to some set Σ of modal formulas, has not been studied here. Can one prove that such an α will be equivalent to a single modal formula already?

7.1 *Theorem.* $LMLLp \rightarrow MMLMp \in \bar{M}1 - M1$.

Proof. That $LMLLp \rightarrow MMLMp \in \bar{M}1$, is easily seen by noting that $\bar{E}(LMLLp \rightarrow MMLMp, \forall x \exists y Rxy)$. The implication from right to left is obvious: if each world in W has an R-successor, then $L\varphi$ implies $M\varphi$ for all φ. For the other direction, suppose that some $w \in W$ has no R-successor: any formula $L\varphi$ will the be true at w and any formula $M\varphi$ will be false. In other words, $LMLLp \rightarrow MMLMp$ is automatically falsified at w.

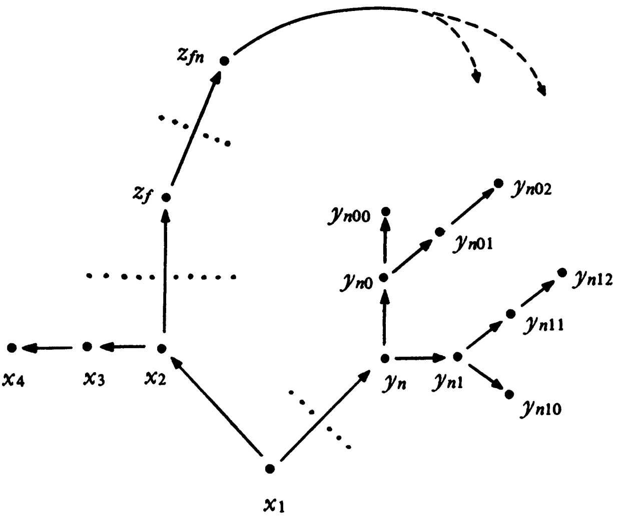
To prove that $LMLLp \rightarrow MMLMp \notin M1$, a frame $\mathsf{F} = \langle W, R\rangle$ and $w \in W$ will be presented such that $\mathsf{F} \models LMLLp \rightarrow MMLMp[w]$; whereas, for no countable elementary subframe F' of F (containing some specified countable subset of W), $\mathsf{F}' \models LMLLp \rightarrow MMLMp[w]$. By the Löwenheim-Skolem Theorem for L_0, this means that it cannot be equivalent to an L_0-formula.

Let $W = \{x_1, x_2, x_3, x_4\} \cup \{y_n, y_{ni}, y_{nij} \mid n \in IN, i \in \{0, 1\}, j \in \{0, 1, 2\}\} \cup \{z_f, z_{fn} \mid n \in IN, f\colon IN \rightarrow \{0, 1\}\}$.

(The notational convention involved here is that "$\{x_i, y_j \mid i \in I, j \in J\}$" is short for "$\{x_i \mid i \in I\} \cup \{y_j \mid j \in J\}$", and similarly for longer sequences x_i, y_j, z_k,... and sequences in which double or triple subscripts are used.)

Let $R = \{\langle x_1, x_2\rangle, \langle x_2, x_3\rangle, \langle x_3, x_4\rangle \cup \{\langle x_2, z_f\rangle, \langle z_f, z_{fn}\rangle, \langle z_{fn}, y_{nf(n)2}\rangle \mid f\colon IN \rightarrow \{0, 1\}, n \in IN\} \cup \{\langle x_1, y_n\rangle, \langle y_n, y_{ni}\rangle, \langle y_{ni}, y_{nij}\rangle, \langle y_{ni1}, y_{ni2}\rangle \mid n \in IN, i \in \{0, 1\}, j \in \{0, 1\}\}$.

Let V be any valuation on F satisfying $\langle \mathsf{F}, V\rangle \models LMLLp[x_1]$. We show that $\langle \mathsf{F}, V\rangle \models MMLMp[x_1]$; thereby establishing that $\mathsf{F} \models LMLLp \rightarrow MMLMp[x_1]$. Since $\langle \mathsf{F}, V\rangle \models LMLLp[x_1]$, $\langle \mathsf{F}, V\rangle \models MLLp[y_n]$ for all $n \in IN$. So, for all $n \in IN$, either $\langle \mathsf{F}, V\rangle \models LLp[y_{n0}]$, in which case $\langle \mathsf{F}, V\rangle \models p[y_{n02}]$, or $\langle \mathsf{F}, V\rangle \models LLp[y_{n1}]$, in which case $\langle \mathsf{F}, V\rangle \models p\,[y_{n12}]$. Let $f\colon IN \rightarrow \{0, 1\}$ satisfy $\langle \mathsf{F}, V\rangle \models p\,[y_{nf(n)2}]$ for all $n \in IN$. Then $\langle \mathsf{F}, V\rangle \models Mp\,[z_{fn}]$



for all $n \in IN$, so $\langle \mathbf{F}, V\rangle \models LMp[z_f]$, and therefore $\langle \mathbf{F}, V\rangle \models MMLMp[x_1]$.

Now let $\mathbf{F}'$ be any countable elementary subframe of $\mathbf{F}$ with a domain containing $\{x_1, x_2, x_3, x_4\} \cup \{y_n, y_{ni}, y_{nij} \mid n \in IN, i \in \{0, 1\}, j \in \{0, 1, 2\}\}$. Take any $z_f \in W—W'$ and put $V(p) = \{y_{nf(n)2} \mid n \in IN\}$. Then $\langle \mathbf{F}', V\rangle \models LMLLp[x_1]$, because $\langle \mathbf{F}', V\rangle \models Lp[x_4]$, $\langle \mathbf{F}', V\rangle \models LLp\,[x_3]$, $\langle \mathbf{F}', V\rangle \models MLLp\,[x_2]$, and also $\langle \mathbf{F}', V\rangle \models p[y_{nf(n)2}]$, $\langle \mathbf{F}', V\rangle \models Lp[y_{nf(n)1}]$, $\langle \mathbf{F}', V\rangle \models Lp[y_{nf(n)0}]$, $\langle \mathbf{F}', V\rangle \models LLp[y_{nf(n)}]$, $\langle \mathbf{F}', V\rangle \models MLLp[y_n]$. Moreover, $\langle \mathbf{F}', V\rangle \not\models MMLMp[x_1]$; for $\langle \mathbf{F}', V\rangle \not\models Mp[x_4]$, $\langle \mathbf{F}', V\rangle \not\models LMp[x_3]$, and, for all $i, n \in IN$, $\langle \mathbf{F}', V\rangle \not\models Mp[y_{ni0}]$, $\langle \mathbf{F}', V\rangle \not\models LMp[y_{ni}]$, and finally $\langle \mathbf{F}', V\rangle \not\models LMp[z_g]$ for any $z_g \in W'$. To see this, note that $z_g \neq z_f$, so $g \neq f$ and, for at least one $n \in IN$, $g(n) \neq f(n)$. For such an n, $\langle \mathbf{F}', V\rangle \not\models Mp[z_{gn}]$, since $\langle \mathbf{F}', V\rangle \not\models p[y_{ng(n)2}]$, and therefore $\langle \mathbf{F}', V\rangle \not\models LMp[z_g]$. It follows that $\mathbf{F}' \not\models LMLLp \rightarrow MMLMp[x_1]$. QED.

Yet another example of this phenomenon will be presented now. (The formulas involved are important for later developments as well.) The relevant facts are as follows.

(i) For all F $(= \langle W, R\rangle)$ and $w \in W$,
$\mathsf{F} \models Lp \rightarrow LLp[w] \Leftrightarrow \mathsf{F} \models \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))[w]$

(ii) $LMp \rightarrow MLp$ is not in $M1$ (cf. chapter 10, theorem 2)

(iii) $(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp)$ is in $\bar{M}1$.

For, $Lp \rightarrow LLp$ holds on a frame if and only if its relation is transitive, and we have the following result.

7.2 *Lemma.* For all *transitive* frames F and $w \in W$,

$$\mathsf{F} \models LMp \rightarrow MLp[w] \Leftrightarrow \mathsf{F} \models \exists y(Rxy \wedge \forall z(Ryz \rightarrow z = y))[w].$$

Proof. The direction from right to left is obvious. For the converse, suppose that $\mathsf{F} \models \forall y(Rxy \rightarrow \exists z(Ryz \wedge z \neq y))[w]$, and apply the following result to $TC(\mathsf{F}, w)$. QED.

7.3 *Lemma* (*AC*). For any frame $\mathsf{F} = \langle W, R\rangle$ with R transitive, the following two statements are equivalent:

(a) $\mathsf{F} \models \forall x \exists y(Rxy \wedge x \neq y)$;

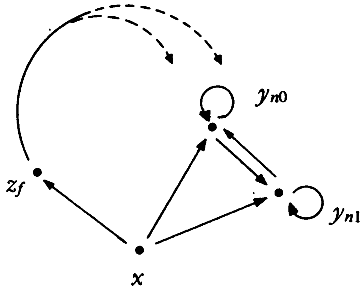
(b) There exists an $X \subseteq W$ such that both X and $W - X$ are *cofinal* in F; i.e., $\forall w \in W \exists v \in X\ Rwv$ and $\forall w \in W \exists v \in (W - X)\ Rwv$.

Proof. Cf. [8]. QED.

7.4 *Theorem.* $(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp) \in \bar{M}1 - M1$.

Proof. That this formula belongs to $\bar{M}1$ is obvious: it is defined by $\forall x \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz)) \wedge \forall x \exists y(Rxy \wedge \forall z(Ryz \rightarrow z = y))$. To see that it is not in $M1$, an argument similar to the one establishing theorem 7.1 is needed. Consider the frame $\mathsf{F} = \langle W, R\rangle$, where

$W = \{x\} \cup \{y_{ni} \mid n \in IN, i \in \{0, 1\}\} \cup \{z_f \mid f\colon IN \rightarrow \{0, 1\}\}$,
$R = \{\langle x, y_{ni}\rangle, \langle y_{ni}, y_{nj}\rangle, \langle x, z_f\rangle, \langle z_f, y_{nf(n)}\rangle \mid i, j \in \{0, 1\}, n \in IN, f\colon IN \rightarrow \{0, 1\}\}$.



F $\models Lp \to LLp[x]$, as is easily checked. Moreover, **F** $\models LMp \to MLp[x]$. For, let V be any valuation on **F** such that $\langle \mathbf{F}, V\rangle \models LMp[x]$. Then, for each n and i, $\langle \mathbf{F}, V\rangle \models Mp[y_{ni}]$, whence $\langle \mathbf{F}, V\rangle \models p[y_{ni}]$ for some $i \in \{0, 1\}$. Let f: $IN \to \{0, 1\}$ be defined by $f(n) = 0$ if $\langle \mathbf{F}, V\rangle \models p[y_{n0}]$ and $f(n) = 1$ otherwise. Clearly, $\langle \mathbf{F}, V\rangle \models Lp[z_f]$ and hence $\langle \mathbf{F}, V\rangle \models MLp[x]$.

Now let **F**′ be any countable elementary subframe of **F** containing x and all y_{ni}. Since there were uncountably many z_f's in W, at least one of them, say z_g, must be missing in **F**′. Define $V(p)$ as $\{y_{ng(n)} \mid n \in IN\}$. Then

(i) $\langle \mathbf{F}', V\rangle \models LMp[x]$, but
(ii) $\langle \mathbf{F}', V\rangle \models \neg MLp[x]$;

whence $(Lp \to LLp) \wedge (LMp \to MLp)$ is falsified at x in **F**′ (and this formula cannot belong to *M1*). To see that (i) holds, note that Mp is true at every y_{ni}, and also at any $z_f \in W'$. For, such an f will equal g for at least one argument n. (Otherwise, f would be the function $1 - g$. But, since it is L_0-expressible that, for every z_f, there exists an element "z_{1-f}" — which property holds in **F** — and since, moreover, **F**′ is an L_0-elementary subframe of **F**, z_g would be in W', contrary to the assumption.) To see that (ii) holds, note that Lp holds at no y_{ni}, but also at no $z_f \in W'$, since any such f differs from g in at least one argument n. QED.

A related formula *is* in *M1*, however:

7.5 *Lemma.* $(Lp \to LLp) \wedge L(Lp \to LLp) \wedge (LMp \to MLp) \in M1$.

Proof. It holds that $E(L(Lp \to LLp), \forall y(Rxy \to \forall z(Ryz \to \forall u(Rzu \to Ryu))))$. Now, if **F** $\models (Lp \to LLp) \wedge L(Lp \to LLp)\ [w]$, then $TC(\mathbf{F}, w)$ is transitive. If, in addition, **F** $\models LMp \to MLp\ [w]$ (and hence $TC(\mathbf{F}, w)$

$\models LMp \rightarrow MLp\ [w]$), then, by lemma 7.2,
$TC(\mathbf{F}, w) \models \exists y(Rxy \wedge \forall z(Ryz \rightarrow z = y))\ [w]$, whence
$\mathbf{F} \models \exists y(Rxy \wedge \forall z(Ryz \rightarrow z = y))\ [w]$.
Therefore, $\varphi = (Lp \rightarrow LLp) \wedge L(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp)$ implies $\alpha = \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz)) \wedge \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow \forall u(Rzu \rightarrow Ryu))) \wedge \exists y(Rxy \wedge \forall z(Ryz \rightarrow z = y))$. The converse implication is even easier to prove. QED.

Results like the above are not as innocent as they seem. They depend on the Axiom of Choice (via lemma 7.3) and, conversely, they imply non-constructive principles which are not provable in Zermelo-Fraenkel Set Theory (ZF).

7.6 *Lemma.* For the above φ and α, $E(\varphi,\alpha)$ implies the Axiom of Choice for unordered pairs.

Proof. Let $\{A_i \mid i \in I\}$ be a set of disjoint unordered pairs. The following application of $E(\varphi,\alpha)$ yields a set of representatives for $\{A_i \mid i \in I\}$. Take some w outside $\bigcup_{i \in I} A_i$, and let $R = \{\langle x,y \rangle \mid (x = w \;\&\; y \in \bigcup_{i \in I} A_i)$ or, for some $i \in I$, $x \in A_i \;\&\; y \in A_i\}$, setting $\mathbf{F} = \langle \bigcup_{i \in I} A_i \cup \{w\}, R \rangle$.

$\mathbf{F} \models \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))\ [w]$ and
$\mathbf{F} \models \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow \forall u(Rzu \rightarrow Ryu)))\ [w]$, so
$\mathbf{F} \models (Lp \rightarrow LLp) \wedge L(Lp \rightarrow LLp)\ [w]$.

As $\mathbf{F} \not\models \exists y(Rxy \wedge \forall z(Ryz \rightarrow z = y))\ [w]$, $\mathbf{F} \not\models \varphi\ [w]$, and this can only be the case because $\mathbf{F} \not\models LMp \rightarrow MLp\ [w]$. Now, if V is any valuation on $\mathbf{F}$ for which $\langle \mathbf{F}, V \rangle \models LMp\ [w]$ and $\langle \mathbf{F}, V \rangle \not\models MLp\ [w]$(i.e., $\langle \mathbf{F}, V \rangle \models LM\neg p[w]$), then $V(p) - \{w\}$ is the required set, having exactly one member in common with each A_i. QED.

It was shown in Jech [37] that the Axiom of Choice for unordered pairs is not provable in ZF. Another of Jech's results is needed for an even stronger theorem.

7.7 *Lemma* (Jech [37], p. 96, problem 15). It is not provable in ZF that any linear ordering $\mathbf{F} = \langle W, R \rangle$ without a last element has a subset $X \subset W$ such that both X and $W - X$ are cofinal in $\mathbf{F}$.

We are now ready to prove

7.8 *Theorem.* It is not provable in *ZF* that $(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp)$ belongs to *M1*.

Proof. Suppose this fact could be proven within *ZF*. From this point, one can argue within that theory, and show that the principle mentioned in lemma 7.7 would be provable as well. Let $\mathbf{F} = \langle W,R \rangle$ be any linear ordering without a last element, such that, for no $X \subset W$, both X and $W - X$ are cofinal in $\mathbf{F}$. Then $\mathbf{F} \models LMp \rightarrow MLp$, and, since R is transitive, $\mathbf{F} \models (Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp)$. Let the L_0-sentence β express that R is a linear order without a last element. Let α be an $L\acute{E}$-sentence equivalent to our modal formula. Clearly, $\mathbf{F} \models \beta \wedge \alpha$, so $\beta \wedge \alpha$ is a satisfiable L_0-sentence, which has a *countable* model $\mathbf{F}_1$.

In the proofs of previous theorems (e.g., 7.1), a strong version of the Löwenheim-Skolem theorem was used, stating the existence of suitable countable elementary substructures for infinite frames. Here, a weak version suffices, viz. that every L_0-sentence with a model has a countable model. The latter statement is provable within *ZF* already! One simply constructs a countable model using Beth-tableaus or the Henkin-technique. Note that, in doing so, no appeal is needed to non-constructive principles, such as König's Lemma or the Prime Ideal Theorem; since we are dealing with a countable language, whose formulas may be well-ordered constructively. (Text-book authors often use such non-constructive principles even in the countable case, for their elegance; but, strictly speaking, such a presentation is misleading. Cf. [5], "stelling" 3.) Thus, $\mathbf{F}_1$ is a countable linear ordering without a last element in which $LMp \rightarrow MLp$ holds. But this is a contradiction; for, in such countable orderings, an $X \subset W_1$ is *definable* (using some enumeration of W_1) such that both X and $W_1 - X$ are cofinal in $\mathbf{F}_1$; which means that $LMp \rightarrow MLp$ can be falsified. QED.

Some terminological explanation may be helpful here. *ZF* is strong enough to provide the means for discussing the semantics of modal logic. It is obvious how meta-talk about frames is to be formalized, and the same holds for the modal language. Single modal formulas may receive set-theoretic "proper names" through the use of some "Gödel code", and these will lie within a naturally defined class of "modal formulas".

Notions such as membership of $M1$, etc., then receive an obvious derived set-theoretic formulation.

Sometimes, a purely arithmetical formulation is preferable to a set-theoretic one. E.g., when a set Σ of modal formulas is said to be "arithmetically undefinable", this should be taken to mean that, for no arithmetical formula $\alpha(x)$, the equivalence $IN \models \alpha(\ulcorner\varphi\urcorner) \Leftrightarrow \varphi \in \Sigma$ holds for all modal formulas φ. ($\ulcorner\varphi\urcorner$ is then some *numerical* Gödel code of φ.) But, of course, all this may be expressed within the set-theoretic language too. In this case, the usual convention is to call those ZF-formulas "arithmetical" which have all their quantifiers relativized to V_ω (i.e., the first infinite layer in the cumulative hierarchy; which is rather like IN). Note that the "proper names" of modal formulas referred to above will belong to V_ω. Then, e.g., a set Σ of modal formulas which is definable in ZF could be said to be "provably arithmetical in ZF" if there exists an arithmetical ZF-formula $\alpha(x)$ such that $ZF \vdash$ "for all modal formulas φ, $\varphi \in \Sigma$ iff $\alpha(\varphi)$".

7.9 *Corollary*. $\bar{M}1$ is not provably arithmetical in ZF.

Proof. Suppose that, for some arithmetical ZF-formula $\alpha(x)$, $ZF \vdash$ "for all modal formulas φ, $\varphi \in \bar{M}1$ iff $\alpha(\varphi)$". Since $ZF + AC \vdash$ "$(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp) \in \bar{M}1$" (by the above), it follows that the arithmetical (!) statement α ("$(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp)$") is provable in $ZF + AC$. Now, $ZF + AC$ is *conservative* over ZF with respect to arithmetical statements, and, therefore, $ZF \vdash \alpha($"$(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp)$"$)$. But then, $ZF \vdash$ "$(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp) \in \bar{M}1$"; contradicting theorem 7.8. **QED**.

A similar result may be proven for $M1$. The restriction to "provably arithmetical" makes such results rather weak, however. A more satisfactory answer exists for the case of universal second-order sentences in general. Call such sentences *first-order definable* if a first-order sentence exists (containing only first-order parameters from that second-order sentence) which is logically equivalent to it on the class of all standard models. In chapter 17, the class of first-order definable universal second-order sentences is shown to be arithmetically undefinable.

CHAPTER VIII

A MODEL-THEORETIC CHARACTERIZATION OF FIRST-ORDER DEFINABILITY

A class of frames is called *L_0-elementary* if it is defined by a single L_0-sentence, *L_0-Δ-elementary* if it is defined by a (possibly infinite) set of L_0-sentences, *L_0-Σ-elementary* if it is a union of L_0-elementary classes, and *L_0-$\Sigma\Delta$-elementary* if it is a union of L_0-Δ-elementary classes of frames. As is well-known, a class of frames is L_0-$\Sigma\Delta$-elementary if and only if it is closed under L_0-elementary equivalence. Therefore, the hierarchy does not extend beyond L_0-$\Sigma\Delta$-elementary classes. Thus one gets the following picture:

$$\begin{tikzcd}
& \text{elementary} \arrow[dl] \arrow[dr] & \\
\Sigma\text{-elementary} \arrow[dr] & & \Delta\text{-elementary} \arrow[dl] \\
& \Sigma\Delta\text{-elementary} \arrow[d] & \\
& \text{essentially higher-order} &
\end{tikzcd}$$

For classes of frames defined by universal second-order sentences, the situation is simplified. If an L_0-Δ-elementary class is defined by such a sentence, then it is L_0-elementary, as an easy compactness argument shows. Similarly, L_0-$\Sigma\Delta$-elementary classes collapse to L_0-Σ-elementary ones:

$$\begin{array}{c}\text{elementary}\\ \downarrow\\ \Sigma\text{-elementary}\\ \downarrow\\ \text{essentially higher-order}\end{array}$$

For modal formulas, considered as a special kind of universal second-order sentences (cf. chapter 3), there is a further reduction: any L_0-Σ-elementary class of frames which is defined by a modal formula is L_0-elementary. Thus, only two possibilities remain: a modal formula is either L_0-elementary or not L_0-$\Sigma\Delta$-elementary. For classes of frames defined by *sets* of modal formulas, an additional possibility reappears: these may be L_0-Δ-elementary without being L_0-elementary. We now proceed to formal results.

8.1 *Lemma* (R.I.Goldblatt). Any ultraproduct $\Pi_U \mathbf{F}_i$ of a family of frames $\{\mathbf{F}_i \mid i \in I\}$ is isomorphic to a generated subframe of the ultrapower $\Pi_U \Sigma\{\mathbf{F}_i \mid i \in I\}$.

Proof. The obvious map sending f_U to $(\langle i, f(i)\rangle_{i \in I})_U$ is the required isomorphism. QED.

8.2 *Corollary*. Any class of frames which is closed under generated subframes, disjoint unions, isomorphic images and ultrapowers is closed under ultraproducts.

8.3 *Corollary*. Any L_0-$\Sigma\Delta$-elementary class of frames which is closed under generated subframes and disjoint unions is L_0-Δ-elementary.

Proof. Let $\mathbf{K}$ be such a class. Since $\mathbf{K}$ is L_0-$\Sigma\Delta$-elementary, it is closed under L_0-elementary equivalence and hence under isomorphic images and ultrapowers. (By the Theorem of Łoś, ultrapowers are L_0-elementarily equivalent to their base structures.) Therefore, **K** is closed under ultraproducts, and, by a well-known model-theoretic result (cf. Chang & Keisler [17]), any class of frames which is closed under L_0-elementary equivalence and ultraproducts is L_0-Δ-elementary. QED.

8.4 *Corollary*. For any set Σ of modal formulas, if $FR(\Sigma)$ is L_0-$\Sigma\Delta$-elementary, then it is L_0-Δ-elementary.

8.5 *Corollary*. For any modal formula φ, if $FR(\varphi)$ is L_0-$\Sigma\Delta$-elementary, then it is L_0-elementary.

Proof. Use 8.4 and the fact that, for universal second-order sentences, being L_0-Δ-elementary implies being L_0-elementary. QED.

8.6 *Theorem*. A modal formula belongs to $\bar{M}1$ if and only if it is preserved under L_0-elementary equivalence if and only if it is preserved under ultrapowers.

Proof. If $\varphi \in \bar{M}1$, then it is, clearly, preserved under L_0-elementary equivalence. If φ is preserved under L_0-elementary equivalence, then it is preserved under ultrapowers. Finally, to close the circle, if φ is preserved under ultrapowers, then, by 8.2, $FR(\varphi)$ is closed under ultraproducts. Moreover, the complement of $FR(\varphi)$, being defined by an existential second-order sentence (equivalent to the negation of $\forall x \forall P_1 \dots \forall P_n ST(\varphi)$), is closed under ultraproducts (cf. [17], theorem 4.1.14). Both $FR(\varphi)$ and its complement are closed under isomorphic images, and so, by Keisler's well-known characterization of elementary classes, $FR(\varphi)$ is L_0-elementary. QED.

Theorem 8.6 does not hold for *all* universal second-order sentences, witness the example of (Dedekind) finiteness; which is preserved under ultrapowers, but which is not elementary. In fact, the situation as regards universal second-order sentences is as follows. By lemma 3.12, all such sentences of the form $\forall P_1 \dots \forall P_n \forall x_1 \dots \forall x_m \varphi$, with φ a Boolean combination of L_0-formulas and atomic formulas in $P_1, \dots, P_n$, are preserved under ultraproducts. In chapter 17, it will be shown that those of the form $\forall P_1 \dots \forall P_n \exists x_1 \dots \exists x_m \varphi$, φ as before, are preserved under ultrapowers; but not necessarily under ultraproducts. Finally, any universal second-order sentence with R as its only first-order predicate parameter is logically equivalent to one of the form $\forall P_1 \ldots \forall P_n \; x_1 \ldots \; x_m \forall y_1 \ldots \forall y_s \; \varphi$, φ again as before; so the above results are the best possible ones. Clearly then, such preservation results do not yield much *syntactic* information about which modal formulas are first-order definable. For that purpose, suitable syntactic methods have to be developed (cf. chapter 9).

For the case of *local* definability, the above results turn out to hold in a slightly more complicated version.

8.7 *Theorem.* A modal formula φ is in $M1$ iff, for all frames $\mathbf{F}$ and sets I such that $w_i \in W$ for each $i \in I$, and for all ultrafilters U on I,

$$\forall i \in I \;\; \mathbf{F} \models \varphi[w_i] \text{ only if } \Pi_U \; \mathbf{F} \models \varphi[(\langle w_i \rangle_{i \in I})_U].$$

Proof. If φ is in $M1$, say it is locally equivalent to α, then $\forall i \in I \; \mathbf{F} \models \varphi[w_i]$ implies that $\forall i \in I \; \mathbf{F} \models \alpha[w_i]$. Then, by Łoś' Theorem, $\Pi_U \; \mathbf{F} \models \alpha[(\langle w_i \rangle_{i \in I})_U]$, and hence φ is true at that world as well.

For the converse direction, add an individual constant c to L_0; which yields $L_0(c)$. Consider the class $\mathbf{K}$ of all $L_0(c)$-structures $\langle \mathbf{F}, w \rangle$ which are models of φ in the sense that $\mathbf{F} \models \varphi[w]$.

(i) $\mathbf{K}$ is closed under ultraproducts. For, consider any set $\{\langle \mathbf{F}_i, w_i \rangle \mid i \in I\} \subseteq \mathbf{K}$, and any ultrafilter U on I. Define $\mathbf{F}'$ to be the frame $\Sigma \; \{\mathbf{F}_i \mid i \in I\}$. Clearly, $\langle \mathbf{F}', w_i \rangle$ is a model of φ, by corollary 2.15. So, by assumption, $\langle \Pi_U \mathbf{F}', (\langle w_i \rangle_{i \in I})_U \rangle$ is a model of φ. So, $\Pi_U \{\langle \mathbf{F}_i, w_i \rangle \mid i \in I\} = \langle \Pi_U \{\mathbf{F}_i \mid i \in I\}, (\langle w_i \rangle_{i \in I})_U \rangle$ is a model of φ too; $\Pi_U \{\mathbf{F}_i \mid i \in I\}$ being a generated subframe of $\Pi_U \mathbf{F}'$ (cf. lemma 11.1).

(ii) The complement of $\mathbf{K}$ is also closed under ultraproducts, by a standard argument about existential second-order sentences like above.

It follows from (i) and (ii) - as well as the obvious closure under isomorphic images - that $\mathbf{K}$ is elementary *in* $L_0(c)$ (again by Keisler's theorem). This yields an $L_0(c)$-sentence $\alpha(c)$ equivalent to φ on the class of structures $\langle \mathbf{F}, w \rangle$; which is just another way of saying that $\alpha(x)$ is a local first-order equivalent for φ on the class of frames. **QED**.

This theorem is used in the only proof we have been able to find for

8.8 *Lemma.* For any modal formula φ, $\varphi \in M1$ if and only if $L\varphi \in M1$.

Proof. The direction from left to right is easy. If $E(\varphi, \alpha)$, where $\alpha = \alpha(x)$, then, for some variable y not occurring in α, $E(L\varphi, \forall x(Ryx \rightarrow \alpha))$; because $\mathbf{F} \models L\varphi[w]$ if and only if, for all v such that Rwv, $\mathbf{F} \models \varphi[v]$.

Next, if $\varphi \notin M1$, then, by theorem 8.7, there are $\mathbf{F} = \langle W,R\rangle$, I, $\{w_i \mid i \in I\}$ and U with, for each $i \in I$, $\mathbf{F} \models \varphi[w_i]$, but $\Pi_U\mathbf{F} \not\models \varphi[(\langle w_i\rangle_{i\in I})_U]$. Take some v outside the domain of $\Pi_U\mathbf{F}$, and let $\mathbf{F}_i$ be the frame $\langle W \cup \{v\}, R \cup \{\langle v,w_i\rangle\}\rangle$. Since $\mathbf{F} \models \varphi[w_i]$, $\mathbf{F}_i \models \varphi\ [w_i]$ and $\mathbf{F}_i \models L\ \varphi[v]$. We show that $\Pi_U\mathbf{F}_i \not\models L\varphi[(\langle v\rangle_{i\in I})_U]$, thereby proving that $L\varphi \notin M1$.

For each $i \in I$, $\mathbf{F}_i \models \forall x(Rx_1x \leftrightarrow x = x_2)\ [v, w_i]$ and, therefore, by the Theorem of Los, $\Pi_U\mathbf{F}_i \models \forall x(Rx_1x \leftrightarrow x = x_2)\ [(\langle v\rangle_{i\in I})_U, (\langle w_i\rangle_{i\in I})_U]$. So, $(\langle v\rangle_{i\in I})_U$ has exactly one R-successor in $\Pi_U\mathbf{F}_i$, viz.$(\langle w_i\rangle_{i\in I})_U$. Clearly, $\mathbf{F} \underset{\rightarrow}{\subseteq} \mathbf{F}_i$, and therefore $\Pi_U\mathbf{F} \underset{\rightarrow}{\subseteq} \Pi_U\mathbf{F}_i$. This is an instance of the following general fact (used, e.g., in the proof of lemma 8.1):

If $\mathbf{F}_i \underset{\rightarrow}{\subseteq} \mathbf{F}'_i$ for all $i \in I$, and U is an ultrafilter on I, then $\Pi_U\mathbf{F}_i \underset{\rightarrow}{\subseteq} \Pi_U\mathbf{F}'_i$.

(The proof of this is straightforward.) Now let V be any valuation on $\Pi_U\mathbf{F}$ such that $\langle \Pi_U\mathbf{F}, V\rangle \models \neg\varphi[(\langle w_i\rangle_{i\in I})_U]$. V is also a valuation on $\Pi_U\mathbf{F}_i$, and, by lemma 2.11 (the "Generation Theorem"), $\langle \Pi_U\ \mathbf{F}_i, V\rangle \models \neg\varphi$ $[(\langle w_i\rangle_{i\in I})_U]$. This implies that $\langle \Pi_U\ \mathbf{F}_i, V\rangle \models \neg L\varphi[(\langle v\rangle_{i\in I})_U]$. QED.

Up to now, only preservation under generated subframes, disjoint unions and elementary equivalence has been used. Here is a model-theoretic result involving the remaining two main notions of chapter 2.

8.9 *Theorem.* If a class of frames is closed under elementary equivalence and p-morphic images, then it is closed under ultrafilter extensions.

Proof. It will be shown that, for any frame $\mathbf{F}$, $ue(\mathbf{F})$ is a p-morphic image of some $\mathbf{F}'$ elementarily equivalent to $\mathbf{F}$. Let $\mathbf{F} = \langle W, R\rangle$. Add, for each $X \subseteq W$, a distinct unary predicate constant c_X to L_0. Expand $\mathbf{F}$ to a structure $\mathbf{F}_1$ for this language in the obvious way. Using a familiar model-theoretic construction, take an elementary extension $\mathbf{M} = \langle \mathbf{F}', \langle c'_X\rangle_{X\subseteq W}\rangle$ of $\mathbf{F}_1$ which is *saturated* with respect to sets of formulas containing at most one parameter from W'. (Note that $\mathbf{F}'$ is elementarily equivalent to $\mathbf{F}$.) The function f defined for $w \in W'$ by $f(w) = \{X \subseteq W \mid w \in c'_X\}$ is a p-morphism from $\mathbf{F}'$ onto $ue(\mathbf{F})$; as will be shown now.

(i) f is *well-defined*; for, $f(w)$ is an ultrafilter on $\mathbf{F}$, in view of the equivalences $\forall z(\neg c_Xz \leftrightarrow c_{W-X}z)$ and $\forall z((c_Xz \wedge c_Yz) \leftrightarrow c_{X\cap Y}z)$, which are true in $\mathbf{F}_1$ and, therefore, in $\mathbf{M}$.

(ii) f is *onto*; since any ultrafilter U on $\mathbf{F}$ corresponds to the finitely satisfiable set $\{c_X z \mid X \in U\}$ on $\mathbf{M}$; which is satisfied by some $w \in W'$, because $\mathbf{M}$ is saturated with respect to such sets.

(iii) If $w, v \in W'$ and $R'wv$, then $R_{\mathbf{F}}f(w)f(v)$. For, it suffices to show that, if $w \in c'_{l(X)}$ (where l is the set-theoretic operation defined in chapter 4), then $v \in c'_X$; and this follows from the truth in $\mathbf{F}_1$ (and, therefore, in $\mathbf{M}$) of $\forall y \forall z((c_{l(X)}y \wedge Ryz) \rightarrow c_X z)$.

(iv) If $R_{\mathbf{F}}f(w)U$, then a $v \in W'$ may be found such that $R'wv$ and $f(v) = U$. Consider $\Sigma = \{c_X z \mid X \in U\} \cup \{Rwz\}$. This set is finitely satisfiable in $\mathbf{M}$. For, if $X_1, ..., X_k \in U$, then $\{c_{X_1} z, ..., c_{X_k} z, Rwz\}$ is satisfiable in $\mathbf{M}$. (To see this, note first that $X = X_1 \cap ... \cap X_k \in U$. Next, suppose that the above set were not satisfiable. Then the following formula is true in $\mathbf{M}$: $\forall y(Rwy \rightarrow \neg c_X y)$ and so $\forall y(Rwy \rightarrow c_{W-X} y)$. Moreover, $\forall z(\forall y(Rzy \rightarrow c_{W-X} y) \rightarrow c_{l(W-X)} z)$ holds in $\mathbf{F}_1$, and, therefore, in $\mathbf{M}$. So, $w \in c'_{l(W-X)}$ and $l(W - X) \in f(w)$. But, by the definition of $R_{\mathbf{F}}$, this means that $W - X \in U$: a contradiction.) Since $\mathbf{M}$ is saturated with respect to sets like Σ, a v as described exists. QED.

This lemma was inspired by Fine [24]. Unfortunately, no converse holds, as is shown by the following example, adapted from that same paper.

8.10 *Lemma.* $ML(p \vee q) \rightarrow M(Lp \vee Lq)$ is preserved under ultrafilter extensions; but this formula is not preserved under elementary equivalence.

Proof. To prove that this formula is preserved under ultrafilter extensions, we show that it is *canonical* in the sense of definition 6.11 (i.e., it is preserved in passing from a descriptive general frame — cf. definition 4.6 — in which it holds to the underlying (full) frame). For, then, if it holds in some frame $\mathbf{F}$ (whence it also holds in $\langle ue(\mathbf{F}), W\rangle$ as defined in definition 4.5), it will hold in $ue(\mathbf{F})$; because $\langle ue(\mathbf{F}), W\rangle$ is a descriptive general frame. Suppose, therefore, that $\langle \mathbf{F}, \mathbf{W}\rangle$ is a descriptive general frame in which $ML(p \vee q) \rightarrow M(Lp \vee Lq)$ holds. It will be proven that $\forall x \forall y(Rxy \rightarrow \exists z(Rxz \wedge \forall u(Rzu \rightarrow (Ryu \wedge \forall v(Rzv \rightarrow u = v)))))$ holds in $\mathbf{F}$; which L_0-sentence implies $ML(p \vee q) \rightarrow M(Lp \vee Lq)$, as is easily checked.

Let $x, y \in W$ such that Rxy. A $z \in W$ is to be found such that (i)

Rxz, (ii) $\forall u(Rzu \rightarrow Ryu)$ and (iii) $\forall u \forall v(Rzu \rightarrow (Rzv \rightarrow u = v))$. Consider $\{X \in \mathbf{W} \mid x \in l(X)\} \cup \{l(Y) \mid Y \in \mathbf{W} \;\&\; y \in l(Y)\} \cup \{l(Z) \cup l(W - Z) \mid Z \in \mathbf{W}\}$. It suffices to show that this set has the finite intersection property; for — in that case — it can be extended to an ultrafilter on $\mathbf{W}$ with a single-point intersection $\{z\}$ such that (i), (ii) and (iii) are satisfied. ($\langle \mathbf{F}, \mathbf{W}\rangle$'s being descriptive is used here repeatedly.) Suppose, for the sake of reductio ad absurdum, that, for some $X_1, ..., X_n, Y_1, ..., Y_m, Z_1, ..., Z_s$, $\bigcap_{i=1}^{n} X_i \cap \bigcap_{j=1}^{m} l(Y_j) \cap \bigcap_{k=1}^{s} (l(Z_k) \cup l(W - Z_k)) = \varnothing$. Let $X = \bigcap_{i=1}^{n} X_i$ and let $Y = \bigcap_{j=1}^{m} Y_j$. Clearly, $x \in l(X)$ $(= \bigcap_{i=1}^{n} l(X_i))$ and $y \in l(Y)$ $(= \bigcap_{j=1}^{m} l(Y_j))$. We have that $X \cap l(Y) \cap \bigcap_{k=1}^{s} (l(Z_k) \cup l(W - Z_k)) = \varnothing$, or $X \subseteq W - (l(Y) \cap \bigcap_{k=1}^{s} (l(Z_k) \cup l(W - Z_k)))$. Therefore, $l(X) \subseteq l(W - (l(Y) \cap \bigcap_{k=1}^{s} (l(Z_k) \cup l(W - Z_k))))$ and — since $x \in l(X)$ — x belongs to the latter set. Now note that the following formula is derivable from $ML(p \vee q) \rightarrow M(Lp \vee Lq)$ in K: $L\neg(Lq \wedge \prod_{k=1}^{s} (Lr_i \vee L\neg r_i)) \rightarrow L\neg Lq$ (*). To see this, assume that $\neg L\neg\, Lq$, i.e., MLq. Clearly, then, $ML((q \wedge r_1) \vee (q \wedge \neg r_1))$ and, therefore, $M(L(q \wedge r_1) \vee L(q \wedge \neg r_1))$ (and so $M(Lq \wedge (Lr_1 \vee L\neg r_1))$). The same argument with respect to r_2 yields $M(L(q \wedge r_1 \wedge r_2) \vee L(q \wedge r_1 \wedge \neg r_2) \vee L(q \wedge \neg r_1 \wedge r_2) \vee L(q \wedge \neg r_1 \wedge \neg r_2))$ (and so $M(Lq \wedge (Lr_1 \vee L\neg r_1) \wedge (Lr_2 \vee L\neg r_2))$. It is obvious that, in the end, one obtains $M(Lq \wedge \prod_{k=1}^{s} (Lr_i \vee L\neg r_i))$. Now, since $\langle \mathbf{F}, \mathbf{W}\rangle \models$ (*), $x \in l(W - l(Y))$, and hence $y \in W - l(Y)$ (Rxy holds!), contradicting the fact that $y \in l(Y)$.

To see that $ML(p \vee q) \rightarrow M(Lp \vee Lq)$ is not preserved under elementary equivalence, consider the following frame $\mathbf{F} = \langle W, R\rangle$. Let U be a free ultrafilter on IN. Define W to be $\{IN\} \cup U \cup IN$ (W is uncountable!) and R as $\{\langle IN, X\rangle \mid X \in U\} \cup \{\langle X, n\rangle \mid X \in U \;\&\; n \in X\} \cup \{\langle n, n\rangle \mid n \in IN\}$. It is easily checked that our formula holds at all worlds different from IN. Now, let V be any valuation on $\mathbf{F}$ such that $\langle \mathbf{F}, V\rangle \models ML(p \vee q)[IN]$. Then, for some $X \in U$, $\langle \mathbf{F}, V\rangle \models L(p \vee q)[X]$; i.e., $X \subseteq V(p \vee q)$, whence $V(p \vee q) \cap IN \in U$. Since $V(p \vee q) \cap IN = (V(p) \cap IN) \cup (V(q) \cap IN)$, one of these sets must belong to U (U being an ultrafilter), say $V(p) \cap IN \in U$. So, $\langle \mathbf{F}, V\rangle \models Lp[V(p) \cap IN]$, whence $\langle \mathbf{F}, V\rangle \models MLp[IN]$.

Now, take any countable L_0-elementary subframe $\mathbf{F}'$ of $\mathbf{F}$ with W'

containing IN and all $n \in IN$. One easily constructs disjoint subsets A, B of IN such that, for all remaining members X of U in W', $X \cap A \neq \emptyset$ and $X \cap B \neq \emptyset$. For, let $X_1, \ldots, X_n, \ldots$ be an enumeration of those members. Define $A_0 = \emptyset$, $B_0 = \emptyset$. Next, suppose that finite disjoint sets A_n, B_n have been defined such that, for each i $(1 \leqslant i \leqslant n)$, $X_i \cap A_n \neq \emptyset$, $X_i \cap B_n \neq \emptyset$. Let k, l be any two distinct elements of X_{n+1} which are not in $A_n \cup B_n$. (Such elements exist since X_{n+1} is infinite, being an element of a free ultrafilter.) Define A_{n+1} to be $A_n \cup \{k\}$ and B_{n+1} as $B_n \cup \{l\}$. $A = \bigcup_n A_n$ and $B = \bigcup_n B_n$ are the required sets. Setting $V(p) = A$ and $V(q) = IN - A$ then falsifies $ML(p \vee q) \rightarrow M(Lp \vee Lq)$ at IN. QED.

CHAPTER IX

THE METHOD OF SUBSTITUTIONS

$M1$ is a more suitable object for syntactic studies than $\bar{M}1$. We will now take a closer look at this set. The following lemmas list some of its simple properties (and of E).

9.1 *Lemma.* For all modal formulas φ and ψ and all L_0-formulas α and β,

$$E(\varphi, \alpha) \ \& \ E(\psi, \beta) \Rightarrow E(\varphi \wedge \psi, \alpha \wedge \beta)$$
$$E(\varphi, \alpha) \ \& \ E(\psi, \beta) \Rightarrow E(\varphi \vee \psi, \alpha \vee \beta),$$

provided that φ and ψ have no proposition letters in common,

$$E(\varphi, \alpha) \Leftrightarrow E([\neg p/p]\varphi, \alpha), \text{ for all proposition letters } p.$$

Proof. For all modal formulas φ and ψ, $\mathbf{F} \models \varphi \wedge \psi[w]$ iff $\mathbf{F} \models \varphi[w]$ and $\mathbf{F} \models \psi[w]$. If φ and ψ have no proposition letters in common, then $\mathbf{F} \models \varphi \vee \psi[w]$ iff $\mathbf{F} \models \varphi[w]$ or $\mathbf{F} \models \psi[w]$. This is provable using lemma 2.4. Finally, lemma 2.5 implies that, for all proposition letters p, $\mathbf{F} \models \varphi[w]$ iff $\mathbf{F} \models [\neg p/p]\varphi[w]$. **QED.**

9.2 *Corollary.* For all modal formulas φ and ψ,

$$\varphi \in M1 \ \& \ \psi \in M1 \Rightarrow \varphi \wedge \psi \in M1$$
$$\varphi \in M1 \ \& \ \psi \in M1 \Rightarrow \varphi \vee \psi \in M1,$$

provided that φ and ψ have no proposition letters in common,

$$\varphi \in M1 \Leftrightarrow [\neg p/p]\varphi \in M1, \text{ for all proposition letters } p$$
$$\varphi \in M1 \Leftrightarrow L\varphi \in M1.$$

Proof. The first three assertions follow from lemma 9.1. The fourth assertion is lemma 8.8. **QED**.

9.3 *Lemma.* The following implications do not hold for all modal formulas φ and ψ,

(i) $\varphi \in M1 \Rightarrow \neg\varphi \in M1$
(ii) $\varphi \in M1 \Rightarrow M\varphi \in M1$
(iii) $\varphi \in M1$ & $\psi \in M1 \Rightarrow (\varphi \rightarrow \psi) \in M1$
(iv) $\varphi \in M1 \Rightarrow [\neg p/q]\varphi \in M1$
(v) $\varphi \wedge \psi \in M1 \Rightarrow \varphi \in M1$ & $\psi \in M1$.

Proof. In chapter 10, the modal formula $LMp \rightarrow MLp$ is shown to be outside of $M1$. This formula is equivalent to $\neg(LMp \wedge LM\neg p)$ and to $M(Mp \rightarrow Lp)$. On the other hand, the following formulas *are* in $M1$: LMp, $LM\neg p$, MLp and $Mp \rightarrow Lp$; with L_0-equivalents $\neg\exists yRxy$, $\neg\exists yRxy$, $\exists y(Rxy \wedge \neg\exists zRyz)$ and $\forall y(Rxy \rightarrow \forall z(Rxz \rightarrow z = y))$, respectively. Thus, (i), (ii) and (iii) are obvious.

For (iv), consider $\varphi = (Mp \wedge Mq) \rightarrow M(p \wedge Mq)$. $\varphi \in M1$, because $E(\varphi, \forall y(Rxy \rightarrow \forall z(Rxz \rightarrow Ryz)))$. But $[\neg p/q]\ \varphi = (Mp \wedge M\neg p) \rightarrow M(p \wedge M\neg p)$ (which is equivalent to $\neg M(p \wedge M\neg p) \rightarrow \neg(Mp \wedge M\neg p)$, i.e., to $L(p \rightarrow Lp) \rightarrow (Mp \rightarrow Lp)$), is not in $M1$. In fact, it is not even in $\bar{M}1$; as will be shown in chapter 10, lemma 1.

(v) follows from lemma 7.5: $(Lp \rightarrow LLp) \wedge L(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp) \in M1$; but $LMp \rightarrow MLp \in M1$. Note that this example also shows that $M1$ is not closed under K-derivability. **QED**.

Recall the following useful notions. $\top$ and $\perp$ were signs for formulas which are everywhere true and everywhere false, respectively. *Closed* formulas were formulas containing no proposition letters (but only these two propositional constants). Two new concepts are the following.

9.4 *Definition.* A modal formula φ is *monotone in the proposition letter p*, if, for all models $\mathbf{M} = \langle W, R, V\rangle$, all $w \in W$ and all valuations V' such that $V'(p) \supseteq V(p)$ (but otherwise the same as V),

$$\mathbf{M} \models \varphi[w] \text{ only if } \langle W, R, V'\rangle \models \varphi[w].$$

9.5 *Definition.* A modal formula is *positive* if it is constructed using only $\top$, $\perp$, proposition letters, $\wedge$, $\vee$, L and M.

Any positive formula is monotone in all its proposition letters. It seems likely that a converse result could be proven using the elementary chain method of chapter 15.

9.6 *Lemma.* Any closed formula is in $M1$. If a modal formula φ is monotone in p, then $\varphi \in M1$ if and only if $[\bot/p]\varphi \in M1$.

Proof. Treating $\top$ and $\bot$ as primitives, we add the following clauses to definition 3.1:

$$ST(\top) = \forall x(Rxx \rightarrow Rxx) \quad \text{and} \quad ST(\bot) = \forall x\neg(Rxx \rightarrow Rxx).$$

Then $ST(\varphi)$ will be an L_0-formula for every closed modal formula φ.

The second assertion is proven by observing that, for any modal formula φ which is monotone in p, and for any frame $\mathbf{F}$ and $w \in W$, $\mathbf{F} \models \varphi[w]$ iff $\mathbf{F} \models [\bot/p]\varphi[w]$. From left to right, this is obvious; and from right to left it follows from the fact that $\{w \in W \mid \mathbf{F} \models \bot[w]\} = \emptyset$ and φ's being monotone in p. QED.

Another quite useful notion from chapter 2 is the *modal degree* $d(\varphi)$ of a formula φ (cf. definition 2.6) which measures the maximum depth of nestings of modal operators in φ. Restricting attention to the modal formulas of degree at most one, i.e., to those formulas in which no iterations of modal operators occur — as described in Lewis [51] — trivializes the problem of characterizing $M1$. This follows from the next result.

9.7 *Lemma.* If a modal formula has degree at most one, then it belongs to $M1$.

Proof. Case 1: $d(\varphi) = 0$. Then no modal operators occur in φ, it is a propositional formula, and there are two possibilities. Either φ is a tautology, in which case $E(\varphi, Rxx \rightarrow Rxx)$, or φ is not a tautology, and $E(\varphi, \neg(Rxx \rightarrow Rxx))$, since a falsifying valuation always exists.

Case 2: $d(\varphi) = 1$. A standard calculation using simple propositional and modal valid equivalences shows that any modal formula of degree 1 whose proposition letters are $p_1, \ldots p_n$ is equivalent to a conjunction of formulas of the form $(P \wedge MQ_1 \wedge \ldots \wedge MQ_k) \rightarrow (MR_1 \vee \ldots \vee MR_s)$; where $P, Q_1, \ldots, Q_k, R_1, \ldots, R_s$ are *state descriptions* of the form $(\neg)p_1 \wedge \ldots \wedge (\neg)p_n$ (with $(\neg)\ p_i = p_i$ or $\neg p_i$). (Extreme cases, in which there are no P,

Q's or R's, are not excluded.) It suffices to give first-order definitions for these implications: the conjunction of these definitions will define the original formula. Assume that no repetitions occur among $Q_1, \ldots, Q_k$ and $R_1, \ldots, R_s$. We distinguish several cases:

(1) Some R_i is some Q_j: α is $x = x$.

(2) No R_i is a Q_j:

(2.1) All 2^n state descriptions are among $R_1, \ldots, R_s$: α is $\exists y Rxy$.

(2.2) Some state description is not among $R_1, \ldots, R_s$:

(2.2.1) P is among $R_1, \ldots, R_s$:

(2.2.1.1) There are no Q_j's: α is Rxx.

(2.2.1.2) There are Q_j's: α is

$$\neg \exists y_1(Rxy_1 \wedge \ldots \wedge \exists y_k(Rxy_k \wedge \prod_{1 \leqslant i \neq j \leqslant k} y_i \neq y_j \wedge \prod_{1 \leqslant i \leqslant k} x \neq y_i \wedge \neg Rxx)\ldots).$$

(2.2.2) P is not among $R_1, \ldots, R_s$:

(2.2.2.1) P is among $Q_1, \ldots, Q_k$: α is

$$\neg \exists y_1(Rxy_1 \wedge \ldots \wedge \exists y_k(Rxy_k \wedge \prod_{1 \leqslant i \neq j \leqslant k} y_i \neq y_j)\ldots).$$

(2.2.2.2) P is not among $Q_1, \ldots, Q_k$:

(2.2.2.2.1) There are no Q_j's: α is $x \neq x$.

(2.2.2.2.2) There are Q_j's: α is

$$\neg \exists y_1(Rxy_1 \wedge \ldots \wedge \exists y_k(Rxy_k \wedge \prod_{1 \leqslant i \neq j \leqslant k} y_i \neq y_j \wedge \prod_{1 \leqslant i \leqslant k} x \neq y_i)\ldots).$$

This completes the list of possibilities: which is, at the same time, the list of first-order relational properties definable by means of modal formulas of degree 1. The proof that the relevant equivalences hold is too tedious to be given here. Going through an example will convince the reader. QED.

In [13] it was also established that modal formulas of degree at most one are *complete* with respect to the properties listed here (cf. definition 6.7).

Lemma 9.7 may also be proven using the characterization of $M1$ obtained in theorem 8.7. The idea is to use the fact that, if the formula in question can be falsified in the ultraproduct, this is due to the existence of "enough" distinct R-successors of $(\langle w_i \rangle_{i \in I})_U$. And this fact can be transferred to $\mathbf{F}$ itself, by the theorem of Łoś.

In order to prove some wider-ranging results, a new method has to be introduced now, called the *method of substitutions*. The idea is roughly the following. For certain modal formulas $\varphi = \varphi(p_1, ..., p_n)$, there exist L_0-formulas $\sigma_1, ..., \sigma_n$ such that, whenever $\langle \mathbf{F}, V\rangle \models \neg\varphi[w]$, then $\mathbf{F} \models \neg[\sigma_1/P_1, ..., \sigma_n/P_n]ST(\varphi)[w]$ (with substitution of the formulas σ_i for the unary predicate constants P_i defined in some suitable fashion.

In other words, the L_0-formula $\varphi^* = [\sigma_1/P_1, ..., \sigma_n/P_n]ST(\varphi)$ implies φ. On the other hand, φ always implies its L_0-substitution instances φ^*; so φ will be *equivalent* to φ^*. Moreover, from modal formulas φ of certain forms to be specified below, such equivalents may be constructed effectively.

In practice, two complications may arise. First, there may be *sets* of possible substitutions (rather than a single one) — say $\sigma_1^1, ..., \sigma_n^1, ..., \sigma_1^k, ..., \sigma_n^k$ — and φ becomes equivalent to a *conjunction* of its L_0-substitution instances. Secondly, possible "parameters" from the domain W may play a role; which are introduced by universal quantifiers in front of $ST(\varphi)$. Their treatment will be apparent from the translation procedures to be presented below.

Before proceeding to the actual results, some useful notations will be introduced:

$M^i\varphi$ abbreviates M ... (i times) ... $M\varphi$
$L^i\varphi$: similarly.

(The case where $i = 0$ is included: e.g., $M^0\varphi = \varphi$.)

R^0xy stands for $x = y$
$R^{n+1}xy$ stands for $\exists z_{n+1}(R^nxz_{n+1} \wedge Rz_{n+1}y)$.

(Here, it is more convenient to think of R^1xy, not as $\exists z_1(x = z_1 \wedge Rz_1y)$, but as Rxy.)

9.8 *Theorem.* If the modal formula ψ is positive and the modal formula φ is constructed using L^ip (for proposition letters p, and $i \in IN$), $\top$, $\bot$, $\vee$ $\wedge$ and M, then $\varphi \rightarrow \psi \in M1$.

Proof. First, one reduces the assertion to be proved to the case without mention of "$\vee$". Use obvious propositional and modal equivalences to rewrite φ as a disjunction of formulas constructed using only formulas of the forms L^ip, $\top$, $\bot$, $\wedge$ and M. Then rewrite $\varphi \rightarrow \psi$ as a conjunction of implications, each of which has one of these disjuncts as its antecedent formula.

Lemma 9.6 helps in removing the proposition letters which occur in $\varphi \rightarrow \psi$, but not in both φ and ψ. (In a sense, these do not contribute anything vital to the formula.) Let p be such a proposition letter. If it occurs in ψ, then $\varphi \rightarrow \psi$ is monotone in p, and $\perp$ may be substituted for it. If it occurs in φ, then $\top$ may be substituted for it. For, by corollary 9.2, $[\neg p/p](\varphi \rightarrow \psi)$ may be considered instead of $\varphi \rightarrow \psi$ itself; and this formula is monotone in p. (And substituting $\perp$ for p in $[\neg p/p](\varphi \rightarrow \psi)$ has the same effect as substituting $\top$ for p in $\varphi \rightarrow \psi$.)

Consider some formula $\varphi \rightarrow \psi$ obtained through these manipulations. Write $ST(\varphi \rightarrow \psi)$ in such a way that no two quantifiers have the same bound variable. In this way, there corresponds, to each occurrence of L and M in $\varphi \rightarrow \psi$, a unique bound variable in $ST(\varphi \rightarrow \psi)$. From $ST(\varphi \rightarrow \psi)$, L_0-formulas $CV(p, \varphi)$ will be extracted for each proposition letter p, which, on substitution in a slightly modified form of $ST(\varphi \rightarrow \psi)$, will yield the required L_0-equivalent.

Consider $ST(\varphi)$ occurring as the antecedent formula in $ST(\varphi \rightarrow \psi)$. Move all existential quantifiers corresponding to occurrences of M in φ to the front. This is possible using the operations that bring formulas into a prenex normal form; because only occurrences of $\wedge$ have to be "crossed". This yields $\exists y_1 \dots \exists y_k \varphi'$; so $ST(\varphi \rightarrow \psi)$ may now be written as $\forall y_1 \dots \forall y_k(\varphi' \rightarrow ST(\psi))$.

Fix a variable u not occurring in $ST(\varphi \rightarrow \psi)$. Let $\bar{p}$ be an occurrence of p in φ. $v(\bar{p})$ is the bound variable y_i in $ST(\varphi)$ corresponding to the innermost occurrence of M in φ the scope of which contains $\bar{p}$; or, if no such occurrence of M exists, $v(\bar{p}) = x$. For the greatest number j such that $\bar{p}$ occurs within a subformula of the form $L^j p$, put $CV(\bar{p}, \varphi) = R^j v(\bar{p})u$. $CV(p, \varphi)$ is defined as the inclusive *disjunction* of all formulas $CV(\bar{p}, \varphi)$, where $\bar{p}$ is an occurrence of p in φ. Finally, take alphabetic variants, if necessary, in order to ensure that the formulas $CV(p, \varphi)$ and $\forall y_1 \dots \forall y_k(\varphi' \rightarrow ST(\psi))$ have no bound variables in common.

The L_0-equivalent $s(\varphi \rightarrow \psi)$ of $\varphi \rightarrow \psi$ is obtained by substituting, for each proposition letter p and corresponding unary predicate constant P, and each individual variable z, $[z/u]CV(p, \varphi)$ for Pz in $\forall y_1 \dots \forall y_k(\varphi' \rightarrow ST(\psi))$.

A number of examples illustrating the above procedure will follow this proof; the remainder of which consists in showing that, for all frames

F and all $w \in W$, $\mathbf{F} \models \varphi \to \psi[w]$ if and only if $\mathbf{F} \models s(\varphi \to \psi)[w]$.

One direction is immediate. If $\mathbf{F} \models \varphi \to \psi[w]$, then, for the proposition letters $p_1, ..., p_n$ occurring in $\varphi \to \psi$,

$$\mathbf{F} \models \forall P_1 \ldots \forall P_n\, ST(\varphi \to \psi)[w], \quad \text{and so}$$
$$\mathbf{F} \models \forall P_1 \ldots \forall P_n \forall y_1 \ldots \forall y_k(\varphi' \to ST(\psi))[w], \quad \text{or}$$
$$\mathbf{F} \models \forall y_1 \ldots \forall y_k \forall P_1 \ldots \forall P_n(\varphi' \to ST(\psi))[w].$$

$s(\varphi \to \psi)$ is an instantiation of the latter formula, so $\mathbf{F} \models s(\varphi \to \psi)[w]$. (Compare the remarks to be made following theorem 9.10.)

For the converse, suppose that, for some valuation V, $\langle \mathbf{F}, V\rangle \models \varphi[w]$. It is to be shown that $\langle \mathbf{F}, V\rangle \models \psi[w]$. Now, clearly, $\langle \mathbf{F}, V\rangle \models \exists y_1 \ldots \exists y_k \varphi'[w]$; and so, for some $w_1, \ldots w_k \in W$, $\langle \mathbf{F}, V\rangle \models \varphi'[w, w_1, ..., w_k]$, where w_i is assigned to y_i for each i $(1 \leqslant i \leqslant k)$. The valuation V' is now defined, for each proposition letter p, by setting $V'(p) = \{v \in W \mid \mathbf{F} \models CV(p, \varphi)[w, w_1, ..., w_k, v]\}$; where v is assigned to u.

It can then be shown that $V'(p) \subseteq V(p)$ for all proposition letters p and $\langle \mathbf{F}, V'\rangle \models \varphi'[w, w_1, ..., w_k]$.

A detailed proof of this would yield no additional insights, for these two assertions are obvious consequences of the definition of the formulas $CV(p, \varphi)$.

Substitute the formulas $CV(p, \varphi)$ for the P's in φ': this gives φ''. Since $\langle \mathbf{F}, V'\rangle \models \varphi'[w, w_1, ..., w_k]$, $\mathbf{F} \models \varphi''[w, w_1, ..., w_k]$. From $\mathbf{F} \models s(\varphi \to \psi)[w]$ it then follows that $\mathbf{F} \models \psi'[w, w_1, ..., w_k]$, where ψ' is obtained from $ST(\psi)$ by the same substitution. This amounts to $\langle \mathbf{F}, V'\rangle \models ST(\psi)[w]$ and, therefore, using the facts that $V'(p) \subseteq V(p)$ for all proposition letters p, and that ψ is monotone in all its proposition letters, it is seen that $\langle \mathbf{F}, V\rangle \models ST(\psi)[w]$; i.e., that $\langle \mathbf{F}, V\rangle \models \psi[w]$. QED.

The following seven examples are well-known modal axioms. The modal logic involved is mentioned between parentheses in each case.

(1) $Lp \to p$ (T)

ST: $\forall y(Rxy \to Py) \to Px$

$CV(p, Lp)$: Rxu

s: $\forall y(Rxy \to Rxy) \to Rxx$, or, after simplification, Rxx.

(2) $Lp \to LLp$ $(S4)$

ST: $\forall y(Rxy \to Py) \to \forall z(Rxz \to \forall v(Rzv \to Pv))$
$CV(p, Lp)$: Rxu
s becomes, after a simplification like above, $\forall z(Rxz \to \forall v(Rzv \to Rxv))$.

(3) $p \to LMp$ (B)
ST: $Px \to \forall y(Rxy \to \exists z(Ryz \wedge Pz))$
$CV(p, p)$: $x = u$
s: $x = x \to \forall y(Rxy \to \exists z(Ryz \wedge x = z))$ or, after simplification, $\forall y(Rxy \to Ryx)$.

(4) $MLp \to Lp$ $(S5)$
ST: $\exists y(Rxy \wedge \forall z(Ryz \to Pz)) \to \forall v(Rxv \to Pv)$
$CV(p, MLp)$: Ryu
s: $\forall y((Rxy \wedge \forall z(Ryz \to Ryz)) \to \forall v(Rxv \to Ryv))$ or, after simplification, $\forall y(Rxy \to \forall v(Rxv \to Ryv))$.

(5) $MLp \to LMp$ $(S4.2)$
is treated similarly; which yields, after simplification,
$\forall y(Rxy \to \forall z(Rxz \to \exists v(Rzv \wedge Ryv)))$.

(6) $(MLp \wedge p) \to Lp$ $(S4.4)$
ST: $(\exists y(Rxy \wedge \forall z(Ryz \to Pz)) \wedge Px) \to \forall v(Rxv \to Pv)$
$CV(p, MLp \wedge p)$: $Ryu \vee x = u$
s: $\forall y((Rxy \wedge \forall z(Ryz \to (Ryz \vee x = z)) \wedge (Ryx \vee x = x)) \to \forall v(Rxv \to (Ryv \vee x = v)))$ or, simplified, $\forall y(Rxy \to \forall v(Rxv \to (Ryv \vee x = v)))$.

(7) $L(Lp \to q) \vee L(Lq \to p)$ $(S4.3)$
This formula has to be rewritten first to $M(Lp \wedge \neg q) \to L(M\neg q \vee p)$, and then, using corollary 12.2, to $M(Lp \wedge q) \to L(Mq \vee p)$.
ST: $\exists y(Rxy \wedge \forall z(Ryz \to Pz) \wedge Qy) \to \forall s(Rxs \to (\exists t(Rst \wedge Qt) \vee Ps))$.
$CV(p, M(Lp \wedge q))$: Ryu
$CV(q, M(Lp \wedge q))$: $y = u$
s: $\forall y((Rxy \wedge \forall z(Ryz \to Ryz) \wedge y = y) \to \forall s(Rxs \to (\exists t(Rst \wedge y = t) \vee Rys)))$ or, simplified, $\forall y(Rxy \to \forall s(Rxs \to (Rsy \vee Rys)))$.

A similar procedure yields, for $L((Lp \wedge p) \to q) \vee L(Lq \to p)$, $\forall y(Rxy \to \forall s(Rxs \to (Rsy \vee Rys \vee s = y)))$.

For our next theorem, a definition is needed.

9.9 *Definition.* *Positive* and *negative* occurrences of a proposition letter p in a modal formula are defined inductively according to the clauses

(i) p occurs positively in p
(ii) p does not occur in $\top$ or $\bot$
(iii) a positive (negative) occurrence of p in φ is a negative (positive) occurrence of p in $\neg\varphi$
(iv) a positive (negative) occurrence of p in φ is a negative (positive) occurrence of p in $\varphi \rightarrow \psi$, but a positive (negative) occurrence of p in $\psi \rightarrow \varphi$
(v) a positive occurrence of p in φ is a positive (negative) occurrence of p in $L\varphi$.

From this definition the following derived rule may be obtained:

(vi) a positive (negative) occurrence of p in φ is a positive (negative) occurrence of p in $\varphi \wedge \psi$, $\psi \wedge \varphi$, $\varphi \vee \psi$, $\psi \vee \varphi$ and $M\varphi$.

The next theorem is slightly more general then 9.8. In a different formulation, it appears in Sahlqvist [66].

9.10 *Theorem.* If a modal formula φ is constructed using proposition letters and their negations, $\top$, $\bot$, $\wedge$, $\vee$, L and M, and φ satisfies, for all proposition letters p occurring in it, either (1) no positive occurrence of p is in a subformula of φ of one of the forms $\psi \wedge \chi$ or $L\psi$ within the scope of some M, or (2) no negative occurrence of p is in a subformula of φ of one of the forms $\psi \wedge \chi$ or $L\psi$ within the scope of some M, then $\varphi \in M1$.

Proof. If some proposition letter p occurs only positively in φ, then φ is monotone in p, and, by lemma 9.6, we can consider $[\bot/p]\varphi$ instead. If a proposition letter p occurs only negatively in φ, then it occurs only positively in $[\neg p/p]\varphi$; a formula which may be considered instead of φ, by corollary 9.2. Then we substitute $\bot$ for p. By corollary 9.2 once more, and contracting double negations, we make every remaining proposition letter satisfy the second condition of the theorem.

Rewrite the negation of the formula just obtained as a formula ψ constructed using (negations of) proposition letters, $\top$, $\bot$, $\wedge$, $\vee$, L and M; by the interchange laws $\neg M\chi \leftrightarrow L\neg\chi$, $\neg L\chi \leftrightarrow M\neg\chi$, De Morgan laws and the law of double negation. Now no positive occurrence of a proposition letter in ψ remains in a subformula of ψ of the form $\chi_1 \vee \chi_2$ or $M\chi$ in the scope of some L.

Any subformula $L\chi$ of ψ is equivalent to a conjunction of formulas of the form L^ip and *n-formulas*; i.e., formulas in which no proposition letter occurs positively. This is proven by induction on χ. The cases $\chi = p, \neg p$, $\top$, $\bot$ and $\chi = \chi_1 \wedge \chi_2$ are trivial. If $\chi = \chi_1 \vee \chi_2$ or $\chi = M\chi_1$, then no proposition letter occurs positively in it; since $L\chi$ satisfies the same condition as ψ. Finally, if $\chi = L\,\chi_1$, use the induction hypothesis and the law $L(\chi_1 \wedge \chi_2) \leftrightarrow (L\chi_1 \wedge L\chi_2)$. Transform ψ into ψ' by replacing occurrences of $L\chi$ which do not lie within the scope of another L by equivalents of the kind described here.

A second induction establishes that each subformula χ of ψ' is equivalent to a disjunction of formulas constructed using formulas of the form L^ip, n-formulas, $\wedge$ and M. The cases $\chi = p, \neg p, \top$, $\bot$ and $\chi = \chi_1 \vee \chi_2$ are trivial. If $\chi = M\chi_1$, then use the law $M(\chi_1 \vee \chi_2) \leftrightarrow (M\chi_1 \vee M\chi_2)$; and if $\chi = \chi_1 \wedge \chi_2$, use the propositional distributive laws. Finally, if $\chi = L\chi_1$, then, by the above, it is either an n-formula or it is of the form L^ip. Applying this result to ψ' itself, one obtains a disjunction $\psi'' = \psi_1 \vee \ldots \vee \psi_n$, with the formulas ψ_i constructed as indicated. ψ'' is obtained by rewriting $\neg\varphi$; so φ is equivalent to $\neg\psi_1 \wedge \ldots \wedge \neg\psi_n$. In view of corollary 9.2, it suffices to consider these formulas $\neg\psi_i$.

$ST(\psi_i)$ can be written in the form $\exists y_1 \ldots \exists y_k\, \psi_i'$, as in the proof of theorem 9.9; but now only with respect to those occurrences of M with a positive occurrence of a proposition letter in their scope. For each proposition letter p, $CV(p, \psi_i)$ may be defined like before; and then substituted in $\forall y_1 \ldots \forall y_k \neg\psi_i'$. This yields the required equivalent $s(\neg\psi_i)$; as may be shown in almost the same way as in the previous proof.

Again it is obvious that $\mathbf{F} \models \neg\psi_i\,[w]$ implies that $\mathbf{F} \models s(\neg\psi_i)\,[w]$. For the converse, suppose that $\mathbf{F} \not\models \neg\psi_i\,[w]$. Then, for some valuation V on $\mathbf{F}$, $\langle \mathbf{F}, V\rangle \models \psi_i\,[w]$ and so $\langle \mathbf{F}, V\rangle \models \psi_i'\,[w, w_1, \ldots, w_k]$ for some $w_1, \ldots, w_k \in W$. Defining V' using the formulas $CV(p, \psi_i)$ as before yields that $\langle \mathbf{F}, V'\rangle \models \psi_i'[w, w_1, \ldots, w_k]$, as well as $V'(p) \subseteq V(p)$ for all proposition letters p. (The second assertion is now needed in proving that n-formulas remain true in the transition from V to V'.) From this it follows that $\mathbf{F} \models \psi_i''\,[w, w_1, \ldots, w_k]$; where ψ_i'' is ψ_i' with the formulas $CV(p, \psi_i)$ substituted for the P's. But $s(\neg\psi_i) = \forall y_1 \ldots \forall y_k \neg\psi_i''$ and, therefore, $\mathbf{F} \models s(\neg\psi_i)\,[w]$. QED.

$M(p \wedge LM\neg p) \rightarrow (MLp \vee LL\neg p)$ is a formula which can be treated using theorem 9.10, but not using theorem 9.8. It will be obvious from previous arguments that any modal formula is equivalent to one constructed using proposition letters and their negations, $\top$, $\bot$, $\wedge$, $\vee$, L and M. Applying the relevant laws in the present case yields

$$L(\neg p \vee MLp) \vee MLp \vee LL\neg p,$$

satisfying the second condition of the theorem. Rewriting its negation yields $M(p \wedge LM\neg p) \wedge LM\neg p \wedge MMp$; which is already a ψ_i in the sense of the above proof. (The only n-formula occurring in it is $LM\neg p$.)
$ST(\psi_i) = \exists y(Rxy \wedge Py \wedge \forall z(Ryz \rightarrow \exists v(Rzv \wedge \neg Pv))) \wedge \forall w(Rxw \rightarrow \exists s(Rws \wedge \neg Ps)) \wedge \exists t(Rxt \wedge \exists r(Rtr \wedge Pr))$.
$CV(p, \psi_i) = (y = u \vee r = u)$.
$s(\neg\psi_i)$ becomes, after simplification,

$$\forall y(Rxy \rightarrow \forall t(Rxt \rightarrow \forall r(Rtr \rightarrow (\forall z(Ryz \rightarrow \exists v(Rzv \wedge v \neq y \wedge v \neq r)) \rightarrow \exists w(Rxw \wedge \forall s(Rws \rightarrow (s = y \vee s = r))))))).$$

The idea in the previous proofs has been to consider a modal formula $\varphi = \forall P_1 ... \forall P_n\, \psi(P_1,...,P_n, R)$, rewrite it, with parameters $y_1,...,y_k$ in front, to get $\forall P_1 ... \forall P_n\, \forall y_1 ... \forall y_k\, \psi'$, and then to find L_0-formulas $\sigma_1,..., \sigma_n$ with free variables among $x, y_1,..., y_k$ to be substituted for $P_1,..., P_n$. This yields an L_0-formula $s(\varphi)$ equivalent to φ. Here the direction from φ to $s(\varphi)$ takes care of itself (a universal instantiation has taken place); but the converse requires proof. Assuming that $\langle \mathsf{F}, V\rangle \models \neg\varphi\,[w]$, it is shown that already $\langle \mathsf{F}, V'\rangle \models \neg\varphi\,[w]$; where V' is a valuation defined by the σ_i's. "Pushing" the σ_i's from the valuation into $\neg\varphi$ then yields a counterexample to $s(\varphi)$.

From this point of view, those modal formulas φ are of interest for which $\langle \mathsf{F}, V\rangle \models \varphi\,[w]$ implies $\langle \mathsf{F}, V_1\rangle \models \varphi\,[w]$ or...or $\langle \mathsf{F}, V_m\rangle \models \varphi\,[w]$, where $V_1,...,V_m$ are valuations assigining L_0-definable subsets of W. Most formulas in $M1$ with which we are acquainted fall into this category, even those not covered by theorem 9.10, such as the ones mentioned in the third and fourth clauses of theorem 10.8. Further investigation of this matter has led to some slight extensions of theorem 9.10 (with liberalized restrictions on the occurrences of proposition letters). E.g., $(L(p \leftrightarrow q) \wedge Mq) \rightarrow q$ can be shown to belong to $M1$. The relevant results are not stated here, however; because the gain in generality is more than off-set by an enormous cost in technical complications.

The two definitions given below describe the class of modal formulas amenable to treatment by the method of substitutions.

9.11 *Definition.* Let α be an L_1-formula of the form $\forall x_1 \dots \forall x_k\, \beta$, with $\beta = \beta(P_1, \dots, P_n, x_1, \dots, x_k, x)$ and $\sigma_1, \dots, \sigma_n$ L_0-formulas having no bound variables in common with β, such that, for a fixed variable s not occurring in α, $\sigma_i = \sigma_i(x_1, \dots, x_k, x, s)$ for all i $(1 \leqslant i \leqslant n)$. Then $[\sigma_1/P_1, \dots, \sigma_n/P_n]\, \alpha$ —i.e., α with subformulas of the form $P_i u$ replaced by $[u/s]\, \sigma_i$ — is a *substitution instance* of α; provided that no x_j becomes bound by an occurrence of $\forall x_j$ different from the first such occurrence in α.

(Technicalities like the above are unavoidable when formulating substitution for formulas. The idea is quite simple, however.)

9.12 *Definition.* M_1^{sub} is the set of modal formulas φ $(= \forall P_1 \dots \forall P_n\, ST(\varphi))$ which are logically implied by the set of substitution instances of L_1-formulas logically equivalent to $ST(\varphi)$.

9.13 *Theorem.*
$M_1^{sub} \underset{\neq}{\subset} M1$
M_1^{sub} is recursively enumerable.

This theorem was proven in the original dissertation [5]. Here, a more elegant view of the matter is obtained by the introduction of the following model-theoretic notion.

9.14 *Definition.* M_1^{def} is the set of modal formulas φ preserved in passing from a general frame $\langle \mathbf{F}, W \rangle$ with W containing all subsets of W definable by means of an L_0-formula and finitely many parameters in W, to the underlying frame $\mathbf{F}$. Formally, if $\langle \mathbf{F}, W \rangle$ is a general frame such that, for any L_0-formula $\alpha = \alpha(x, x_1, \dots, x_k)$ and all $w_1, \dots, w_k \in W$, $\{w \in W \mid \mathbf{F} \models \alpha[w, w_1, \dots, w_k]\} \in W$, then if $w \in W$ such that $\langle \mathbf{F}, W \rangle \models \varphi\, [w]$, then $\mathbf{F} \models \varphi\, [w]$.

9.15 *Theorem.*
$M_1^{sub} \subseteq M_1^{def}$
M_1^{def} is recursively enumerable (and this fact is provable in ZF)
$M_1^{def} \underset{\neq}{\subset} M1.$

Proof. To prove the first assertion, let the modal formula φ ($= \forall P_1 \dots \forall P_n ST(\varphi)$) be in M_1^{sub}. Suppose that $\langle \mathbf{F}, \mathbf{W} \rangle \models \forall P_1 \dots \forall P_n ST(\varphi)[w]$, where $\langle \mathbf{F}, \mathbf{W} \rangle$ is a general frame satisfying the condition of definition 9.14. It is to be shown that $\mathbf{F} \models \forall P_1 \dots \forall P_n ST(\varphi)[w]$. To this end, it will be shown that, for any substitution instance β of any α logically equivalent to $ST(\varphi)$, $\mathbf{F} \models \beta[w]$, whence $\mathbf{F} \models \varphi[w]$; since $\varphi \in M_1^{sub}$. Let $\alpha = \forall x_1 \dots \forall x_k \alpha'$, and $\sigma_1, \dots, \sigma_n$ be as in definition 9.11. (Thus, $\sigma_i = \sigma_i(x_1, \dots, x_k, x, s)$ for all i ($1 \leqslant i \leqslant n$).) Let $\beta = [\sigma_1/P_1, \dots, \sigma_n/P_n]\alpha = \forall x_1 \dots \forall x_k[\sigma_1/P_1, \dots, \sigma_n/P_n]\alpha'$. Consider any $w_1, \dots, w_k \in W$: it will now be shown that $\mathbf{F} \models [\sigma_1/P_1, \dots, \sigma_n/P_n]\alpha'[w_1, \dots, w_k, w]$. For each i ($1 \leqslant i \leqslant n$), $A_i =_{def} \{v \in W \mid \mathbf{F} \models \sigma_i[w_1, \dots, w_k, w, v]$; where v is assigned to $s\} \in \mathbf{W}$, and also $\langle \mathbf{F}, \mathbf{W} \rangle \models \forall P_1 \dots \forall P_n \forall x_1 \dots \forall x_k \alpha'[w]$; whence $\langle \mathbf{F}, \mathbf{W} \rangle \models \forall x_1 \dots \forall x_k \forall P_1 \dots \forall P_n \alpha'[w]$, $\langle \mathbf{F}, \mathbf{W} \rangle \models \forall P_1 \dots \forall P_n \alpha'[w_1, \dots, w_k, w]$, and, finally, $\langle \mathbf{F}, A_1, \dots, A_n \rangle \models \alpha'[w_1, \dots, w_k, w]$. But this last assertion implies that $\mathbf{F} \models [\sigma_1/P_1, \dots, \sigma_n/P_n]\alpha'[w_1, \dots, w_k, w]$: which was to be proven.

(Meanwhile, P. Rodenburg has pointed out that also $M_1^{def} \subseteq M_1^{sub}$, by a simple argument. Thus, our two analyses of the method of substitutions coincide.)

To see that M_1^{def} is recursively enumerable, note that $\varphi \in M_1^{def}$ iff, for all general frames $\langle \mathbf{F}, \mathbf{W} \rangle$ with $\mathbf{W}$ consisting of *exactly* the subsets of W definable by means of an L_0-formula and finitely many parameters in W, $\langle \mathbf{F}, \mathbf{W} \rangle \models \varphi[w]$ only if $\mathbf{F} \models \varphi[w]$, for all $w \in W$. Now this is again equivalent to

$$DS(\varphi) \models \forall P_1 \dots \forall P_n ST(\varphi);$$

where $DS(\varphi) =_{def} \{\alpha \mid \alpha$ is an L_0-formula of the form $\forall x_1 \dots \forall x_k[\sigma_1/P_1, \dots, \sigma_n/P_n]ST(\varphi)$; with $\sigma_i = \sigma_i(x_1, \dots, x_k, x, s)$ an L_0-formula substitutable for P_i in $ST(\varphi)\}$. Because $DS(\varphi)$ consists entirely of L_0-formulas, the above implication is equivalent to $DS(\varphi) \models ST(\varphi)$; whence — by compactness for L_1 — it is equivalent to $\alpha' \models ST(\varphi)$ for some finite conjunction σ' of members of $DS(\varphi)$. Clearly, this last equivalent is recursively enumerable; because, for L_1, $\models$ is a recursively axiomatizable notion.

The preceding proof is formalizable in ZF, notwithstanding the use of the compactness theorem. For, the latter result is only needed for the countable language L_1, and compactness for such a language may be established within ZF. (The point is that Lindenbaum's Extension Lemma can be proven without having recourse to non-constructive principles such as the Prime Ideal Theorem, provided that a well-ordering of (the formulas of) the language be given in advance.)

One part of the third assertion, viz. that $M_1^{def} \subseteq M1$, is easy to prove. For, clearly, $\forall P_1 \ldots \forall P_n ST(\varphi)$ implies each $\alpha \in DS(\varphi)$; so, if $\varphi \in M_1^{def}$, then, for the formula α' mentioned above, $\models \forall P_1 \ldots \forall P_n ST(\varphi) \leftrightarrow \alpha'$, and hence $\varphi \in M1$.

The proof that $M_1^{def} \neq M1$ involves an argument about provability of certain formulas within ZF, which has been included here for its intrinsic interest. There is also a simpler model-theoretic proof, however; witness the argument establishing theorem 9.16 below.

It is tempting to try using the result of chapter 7 stating that $M1$ is not provably arithmetical in ZF, together with the above assertion, to show that M_1^{def} and $M1$ cannot be equal. But, such an argument would only show that M_1^{def} and $M1$ are not *provably* equal in ZF. To establish the stronger claim that these classes are actually different, consider the modal formula $\varphi = (Lp \to LLp) \wedge L(Lp \to LLp) \wedge (LMp \to MLp)$ of chapter 7. We have

(1) $ZF + AC \vdash$ "$\varphi \in M1$"
(2) $ZF \nvdash$ "$\varphi \in M1$".

Now suppose that this formula φ is in M_1^{def}. Then, for some α' as described above, $\varphi \models \alpha'$ and also $\alpha' \models \varphi$. The first of these assertions is trivially provable in ZF (because of the form of α'), and so is the second; for $\alpha' \models \varphi$ ($= \forall P_1 ST(\varphi)$) iff $\alpha' \models ST(\varphi)$, and this last equivalent is provable in ZF (because it is logically provable in L_1). We have thus derived

(3) $ZF \vdash$ "$\varphi \in M_1^{def}$",

which contradicts (2); since the proof of the fact that $M_1^{def} \subseteq M1$ can obviously be given within ZF. Hence, $\varphi \in M1 - M_1^{def}$. QED.

In the sequel, we will be interested in global notions again. $\bar{M}_1^{def}$ is defined in the obvious way: just like in definition 9.14, but for the omission of the parameter w. Before the result corresponding to theorem 9.15 is formulated, however, a model-theoretic lemma is needed.

9.16 *Lemma.* The only sets of natural numbers which are L_0-definable in the frame $\langle IN, <\rangle$ with parameters in IN, are the finite and cofinite ones.

Proof. Note that all elements of IN are L_0-definable in $\langle IN, <\rangle$. Therefore, all finite and cofinite subsets of IN are L_0-definable as well. Now, suppose that some infinite subset of IN whose complement is infinite as well were L_0-definable in $\langle IN, <\rangle$; say by $\alpha = \alpha(x, x_1, ..., x_k)$ with parameters $n_1, ..., n_k$. Since the latter numbers can be defined in L_0, one may as well consider a parameter-free formula α. Then both $\forall x \exists y(Rxy \wedge \alpha(y))$ and $\forall x \exists y(Rxy \wedge \neg\alpha(y))$ would be true in this structure. But, this can be refuted using the well-known fact that $\langle IN, <\rangle$ is L_0-elementarily equivalent to the frame $\langle IN, <\rangle \oplus \langle ZZ, <\rangle$ (i.e., the natural numbers followed by a copy of the integers). For, in the latter structure **F**, the following equivalence holds, for all formulas $\beta = \beta(x)$ and any two members r, s of the "tail" of integers: $\mathbf{F} \models \beta(x)[r]$ if and only if $\mathbf{F} \models \beta(x)[s]$. (To see this, consider the L_0-automorphism g of **F** leaving the natural numbers fixed, but mapping all integers t onto $t + (s - r)$.) But, because of this equivalence, $\forall x \exists y(Rxy \wedge \alpha(y))$ and $\forall x \exists y(Rxy \wedge \neg\alpha(y))$ cannot both be true in **F**: which is a contradiction. QED.

9.17 *Theorem.*
$M1^{def} \subseteq \bar{M}1^{def}$
$\bar{M}1^{def}$ is recursively enumerable (and provably so in ZF)
$\bar{M}1^{def} \subsetneq \bar{M}1$.

Proof. All these assertions are obvious or can be proven just like the corresponding ones in the proof of theorem 9.15. Here, a new argument is presented showing that $\bar{M}1^{def} \neq \bar{M}1$, however. As was shown in chapter 7, the formula $(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp) \in \bar{M}1$. This formula also holds in the general frame $\langle IN, <, W\rangle$, where W consists of all finite and cofinite sets of natural numbers. Now, if it belonged to $\bar{M}1^{def}$, then, by lemma 9.16 (which tells us that this general frame is of the kind described in definition 9.14), it would have to hold in $\langle IN, <\rangle$ as well. But, it does not: e.g., setting "$V(p)$ is the set of all odd numbers" defines a falsifying valuation. QED.

Theorems 9.15 and 9.17 neatly delimit the range of the method of substitutions.

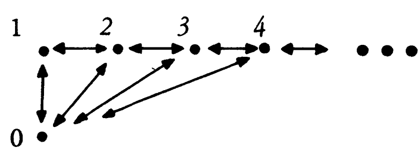
CHAPTER X

DISPROVING FIRST-ORDER DEFINABILITY

Up to now, a few modal formulas have been shown to be outside of $M1$ or $\bar{M}1$. In chapter 3, $L(Lp \rightarrow p) \rightarrow Lp$ turned out to define the transitive frames the converse of whose relation is well-founded. A routine argument involving the compactness theorem establishes that the class of these frames is not L_0-definable. In chapter 7, the Löwenheim-Skolem theorem was used to prove that neither $LMLLp \rightarrow MMLMp$ nor $(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp)$ belong to $M1$. In addition to these two theorems, there are the results of chapter 8 which could be applied: a modal formula is outside of $\bar{M}1$ if it is not preserved under ultrapowers (cf. theorem 8.6) or ultrafilter extensions (cf. theorem 8.9). In fact, from case to case, ad-hoc model-theoretic arguments may be useful, witness the following example (cf. lemma 9.3. (iv)).

10.1 *Lemma.* $L(p \rightarrow Lp) \rightarrow (Mp \rightarrow Lp) \notin \bar{M}1$.

Proof. Consider the frame $\mathbf{F} = \langle W, R \rangle$, where $W = IN$ and $R = \{\langle 0, n \rangle, \langle n, 0 \rangle, \langle n, n+1 \rangle, \langle n+1, n \rangle \mid n = 1, 2, 3, \ldots\}$.

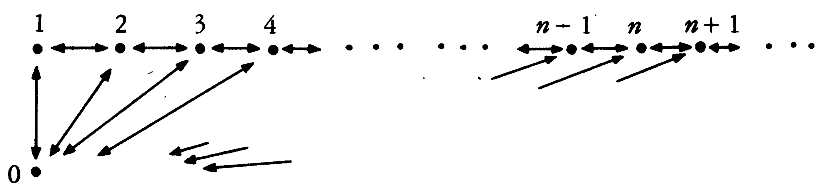


$\mathbf{F} \models L(p \rightarrow Lp) \rightarrow (Mp \rightarrow Lp)$. To see this, consider any valuation V on $\mathbf{F}$

and any $n \in I\!N$ such that $\langle \mathbf{F}, V\rangle \models L(p \to Lp) \wedge Mp[n]$. It is to be shown that $\langle \mathbf{F}, V\rangle \models Lp\ [n]$; i.e., that, for all m such that Rnm, $\langle \mathbf{F}, V\rangle \models p\ [m]$. In case n is 0, $\langle \mathbf{F}, V\rangle \models Mp\ [n]$ implies that $\langle \mathbf{F}, V\rangle \models p\ [k]$ for some $k \geqslant 1$; and $\langle \mathbf{F}, V\rangle \models L(p \to Lp)[0]$ implies that, moreover, $\langle \mathbf{F}, V\rangle \models Lp\ [k]$; whence $\langle \mathbf{F}, V\rangle \models p\ [k+1]$, and $\langle \mathbf{F}, V\rangle \models p\ [k-1]$ (if $k \geqslant 2$). A repetition of this argument shows that $\langle \mathbf{F}, V\rangle \models p\ [m]$ for all $m \geqslant 1$; i.e., that $\langle \mathbf{F}, V\rangle \models Lp\ [0]$.

In case n is greater than 0, $\langle \mathbf{F}, V\rangle \models Mp[n]$ implies that either (i) $\langle \mathbf{F}, V\rangle \models p[0]$ or (ii) $\langle \mathbf{F}, V\rangle \models p[n-1]$ or (iii) $\langle \mathbf{F}, V\rangle \models p\ [n+1]$. Because $\langle \mathbf{F}, V\rangle \models L(p \to Lp)[n]$, (i) $\langle \mathbf{F}, V\rangle \models Lp[0]$ or (ii) $\langle \mathbf{F}, V\rangle \models Lp\ [n-1]$ (and hence $\langle \mathbf{F}, V\rangle \models p\ [0]$ — if $n-1 \geqslant 0$; otherwise, this was already true in the first place — and, therefore, $\langle \mathbf{F}, V\rangle \models Lp\ [0]$) or (iii) $\langle \mathbf{F}, V\rangle \models Lp\ [n+1]$ (and hence $\langle \mathbf{F}, V\rangle \models p\ [0]$ and $\langle \mathbf{F}, V\rangle \models Lp\ [0]$). Thus, in all three cases, both $\langle \mathbf{F}, V\rangle \models Lp\ [0]$ and $\langle \mathbf{F}, V\rangle \models p\ [0]$ hold. It follows easily that $\langle \mathbf{F}, V\rangle \models Lp\ [n]$.

Next, take any proper L_0-elementary extension $\mathbf{F}'$ of $\mathbf{F}$. Some reflection upon the L_0-theory of $\mathbf{F}$ (which is preserved in going from $\mathbf{F}$ to $\mathbf{F}'$) shows that $\mathbf{F}'$ consists of an isomorphic copy of $\mathbf{F}$ together with a number of copies of the integers (under the ordering $\{\langle n, n+1\rangle, \langle n+1, n\rangle \mid n \in Z\!\!Z\}$) attached to 0.



Now $L(p \to Lp) \to (Mp \to Lp)$ is falsifiable in $\mathbf{F}'$ at 0 by setting $V(p) =_{def}$ the worlds belonging to the copy of $\mathbf{F}$ in $\mathbf{F}'$. Lp becomes false at 0, because there are more worlds now than just these; but Mp and $L(p \to Lp)$ are both true.

It follows that our modal formula is not preserved under L_0-elementary extensions: so it cannot be in $\bar{M}1$. QED.

To get some impression of how first-order definability is lost, consider lemma 9.3 (ii). Prefixing of M often leads out of $M1$. Here are a few examples. (Note that the following equivalence is universally valid: $M(\varphi \to \psi) \leftrightarrow (L\varphi \to M\psi)$.)

(i) $Mp \rightarrow Lp \in M1$; for
$E(Mp \rightarrow Lp, \forall y(Rxy \rightarrow \forall z(Rxz \rightarrow z = y)))$.
$M(Mp \rightarrow Lp) \notin M1$; for $LMp \rightarrow MLp \notin M1$ (cf. theorem 10.2 below). On the other hand,

(ii) $Lp \rightarrow Mp \in M1$; for $E(Lp \rightarrow Mp, \exists yRxy)$. But also $M(Lp \rightarrow Mp) \in M1$; for $LLp \rightarrow MMp$ is of the kind described in theorem 9.8: $E(LLp \rightarrow MMp, \exists y(Rxy \wedge \exists zRyz))$.

Two more unproblematic cases are

(iii) $p \rightarrow Lp \in M1$; for $E(p \rightarrow Lp, \forall y(Rxy \rightarrow x = y))$, and also $M(p \rightarrow Lp) \in M1$; since
$E(Lp \rightarrow MLp, \exists y(Rxy \wedge \forall z(Ryz \rightarrow Rxz)))$.

(iv) $Lp \rightarrow p \in M1$: $(E(Lp \rightarrow p, Rxx))$, and also $M(Lp \rightarrow p) \in M1$; because of $E(LLp \rightarrow Mp, \exists y(Rxy \wedge \exists z(Rxz \wedge Rzy))))$.
But conjunction causes trouble:

(v) $(p \rightarrow Lp) \wedge (Lp \rightarrow p)$, i.e., $p \leftrightarrow Lp \in M1$; since $E(p \leftrightarrow Lp, \forall y(Rxy \leftrightarrow x = y)))$, but
$M((p \rightarrow Lp) \wedge (Lp \rightarrow p)) \notin M1$. For, it is equivalent to $M((p \rightarrow (Lp \wedge p)) \wedge (Lp \rightarrow (Lp \wedge p)))$, and so to $M((Lp \vee p) \rightarrow (Lp \wedge p))$, and, therefore, to $L(Lp \vee p) \rightarrow M(Lp \wedge p)$, which is not in $M1$; by theorem 10.3 below.

More systematically, it may be considered if theorem 9.8 is the best possible result, in the sense that violation of its syntactic conditions immediately leads to loss of first-order definability. The theorem describes formulas $\varphi \rightarrow \psi$ with a positive consequent ψ and an antecedent formula φ in which no combinations are allowed of the forms

$L(\ldots M \ldots)$ or $L(\ldots \vee \ldots)$.

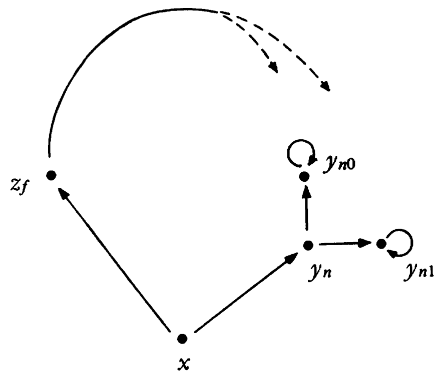
That ψ is to be positive is shown by the example of $L(Lp \rightarrow p) \rightarrow Lp$, which may also be written as $M\neg p \rightarrow M(Lp \wedge \neg p)$, or — by corollary 9.2 — as $Mp \rightarrow M(p \wedge L\neg p)$. The need for the other two requirements is shown by the following negative results.

10.2 *Theorem.* $LMp \rightarrow MLp \notin \bar{M}1$.

Proof. Consider the frame $\mathsf{F} = \langle W, R\rangle$ with $W = \{x\} \cup \{y_n, y_{ni} \mid n \in IN; i \in \{0, 1\}\} \cup \{z_f \mid f\colon IN \rightarrow \{0, 1\}\}$, and $R = \{ \langle x, y_n\rangle, \langle y_n, y_{ni}\rangle, \langle y_{ni}, y_{ni}\rangle \mid n \in IN; i \in \{0, 1\}\} \cup \{ \langle x, z_f\rangle, \langle z_f, y_{nf(n)}\rangle \mid n \in IN; f\colon IN \rightarrow \{0, 1\}\}$.

10.2 *Theorem.* $LMp \to MLp \notin \bar{M}1$.

Proof. Consider the frame $\mathbf{F} = \langle W, R \rangle$ with $W = \{x\} \cup \{y_n, y_{ni} \mid n \in IN; i \in \{0, 1\}\} \cup \{z_f \mid f\colon IN \to \{0, 1\}\}$, and $R = \{ \langle x, y_n \rangle, \langle y_n, y_{ni} \rangle, \langle y_{ni}, y_{ni} \rangle \mid n \in IN; i \in \{0, 1\}\} \cup \{ \langle x, z_f \rangle, \langle z_f, y_{nf(n)} \rangle \mid n \in IN; f\colon IN \to \{0, 1\}\}$.

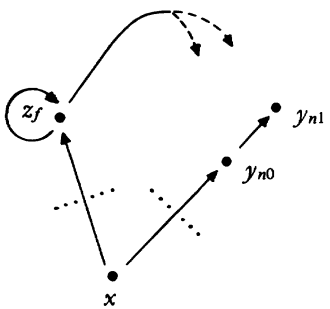


$\mathbf{F} \models LMp \to MLp$. For all w distinct from x, it is easy to see that $\mathbf{F} \models LMp \to MLp[w]$. That $\mathbf{F} \models LMp \to MLp[x]$ is shown as follows. Suppose that $\langle \mathbf{F}, V \rangle \models LMp[x]$. Then, for each $n \in IN$, $\langle \mathbf{F}, V \rangle \models Mp[y_n]$; whence $\langle \mathbf{F}, V \rangle \models p[y_{n0}]$ or $\langle \mathbf{F}, V \rangle \models p[y_{n1}]$. Choose $f : IN \to \{0, 1\}$ such that $\langle \mathbf{F}, V \rangle \models p[y_{nf(n)}]$ for each $n \in IN$. Then, clearly, $\langle \mathbf{F}, V \rangle \models Lp[z_f]$; so $\langle \mathbf{F}, V \rangle \models MLp[x]$.

Now, let $\mathbf{F}'$ be any *countable* L_0-elementary subframe of $\mathbf{F}$ whose domain contains x, y_n, y_{n0}, y_{n1} for all $n \in IN$. There must be some $f : IN \to \{0, 1\}$ such that $z_f \in W\text{-}W'$; because W is uncountable. Setting $V(p) = \{y_{nf(n)} \mid n \in IN\}$ yields a valuation V for which (i) $\langle \mathbf{F}', V \rangle \models LMp[x]$, but (ii) $\langle \mathbf{F}', V \rangle \not\models MLp[x]$. Here, (ii) follows from the fact that Lp holds at no y_n and at no $z_g \in W'$ (because $g \neq f$). To see that (i) holds, note first that Mp holds at each y_n. Moreover, for each $z_g \in W'$, there exists some $n \in IN$ such that $g(n) = f(n)$. (If $g(n) = 1 - f(n)$ for each $n \in IN$, then z_f would be in W'; since the existence of "complementary" worlds z_f is L_0-expressible.) Because of this fact, $\langle \mathbf{F}', V \rangle \models Mp[z_g]$. In other words, $LMp \to MLp$ has been shown to fail (at x) in $\mathbf{F}'$. Hence $LMp \to MLp$ cannot belong to $\bar{M}1$; by the Löwenheim-Skolem theorem for L_0-sentences. QED.

10.3 *Theorem.* $L(Lp \vee p) \rightarrow M(Lp \wedge p) \notin M1$.

Proof. Consider the frame $\mathbf{F} = \langle W, R\rangle$ with $W = \{x\} \cup \{y_{ni} \mid n \in IN; i \in \{0, 1\}\} \cup \{z_f \mid f\colon IN \rightarrow \{0, 1\}\}$, and $R = \{\langle x, y_{n0}\rangle, \langle y_{n0}, y_{n1}\rangle \mid n \in IN\} \cup \{\langle x, z_f\rangle, \langle z_f, z_f\rangle, \langle z_f, y_{nf(n)}\rangle \mid n \in IN; f\colon IN \rightarrow \{0, 1\}\}$.



$\mathbf{F} \models L(Lp \vee p) \rightarrow M(Lp \wedge p)[x]$. For, if $\langle \mathbf{F}, V\rangle \models L(Lp \vee p)[x]$, then $\langle \mathbf{F}, V\rangle \models p\,[y_{n0}]$ or $\langle \mathbf{F}, V\rangle \models p\,[y_{n1}]$ (for each $n \in IN$), and $\langle \mathbf{F}, V\rangle \models p\,[z_f]$ for each world z_f. Then choose any function $f : IN \rightarrow \{0, 1\}$ such that $\langle \mathbf{F}, V\rangle \models p\,[y_{nf(n)}]$ for each $n \in IN$. Clearly, $\langle \mathbf{F}, V\rangle \models Lp \wedge p\,[z_f]$, and hence $\langle \mathbf{F}, V\rangle \models M(Lp \wedge p)\,[x]$.

Then, again, let $\mathbf{F}'$ be any countable L_0-elementary subframe of $\mathbf{F}$ containing x and all worlds y_n. Let $z_f \in W - W'$. Define V by setting $V(p) = \{z_g \mid z_g \in W'\} \cup \{y_{nf(n)} \mid n \in IN\}$. An easy calculation shows that $\langle \mathbf{F}', V\rangle \models L(Lp \vee p)[x]$; but $\langle \mathbf{F}', V\rangle \not\models M(Lp \wedge p)[x]$. So $L(Lp \vee p) \rightarrow M(Lp \wedge p)$ fails at x in $\mathbf{F}'$. QED.

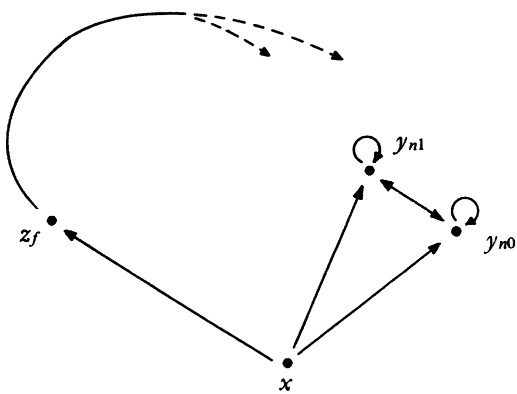
Remark: The formula of 10.3 shows how tricky the present subject is. For, the formula $L(Lp \vee p) \rightarrow MLp$, which seems to violate the conditions of theorem 9.8 in exactly the same way as $L(Lp \vee p) \rightarrow M(Lp \wedge p)$, is in $M1$! For all frames $\mathbf{F}$ and all $w \in W$, $\mathbf{F} \models L(Lp \vee p) \rightarrow MLp\,[w]$ iff $\mathbf{F} \models Lp \rightarrow MLp\,[w]$ (note that $L(Lp \vee p)$ implies $MLp \vee Lp$); and $Lp \rightarrow MLp \in M1$: it was one of the above examples.

Observe also that the statement of theorem 10.3 is weaker than the one made in theorem 10.2. In the proof of 10.2, the relevant formula holds in the frame $\mathbf{F}$ (everywhere, that is), but fails in $\mathbf{F}'$. In the proof of 10.3, however, it holds *at some* w in $\mathbf{F}$, and fails — at that same w — in $\mathbf{F}'$; whence it has no *local* L_0-equivalent. It is often difficult to transform proofs of the latter kind into proofs of the former kind.

For the next result, cf. lemma 8.10.

10.4 *Theorem.* $L(p \vee q) \rightarrow M(Lp \vee Lq) \notin M1$.

Proof. Consider the frame $\mathbf{F} = \langle W, R\rangle$ with $W = \{x\} \cup \{y_{ni} \mid n \in IN; i \in \{0, 1\}\} \cup \{z_f \mid f : IN \rightarrow \{0, 1\}\}$, and $R = \{\langle x, y_{ni}\rangle, \langle y_{ni}, y_{nj}\rangle \mid n \in IN; i, j \in \{0, 1\}\} \cup \{\langle x, z_f\rangle, \langle z_f, y_{nf(n)}\rangle \mid f : IN \rightarrow \{0, 1\}; n \in IN\}$.



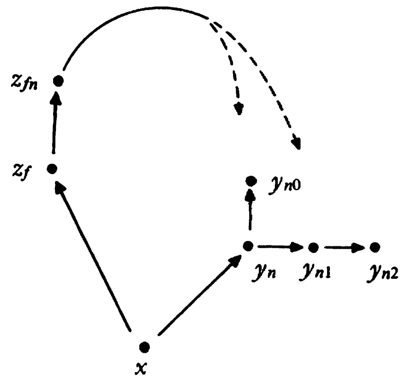
$\mathbf{F} \models L(p \vee q) \rightarrow M(Lp \vee Lq)[x]$. For, if $\langle \mathbf{F}, V\rangle \models L(p \vee q)[x]$, then either, for some $n \in IN$, $\langle \mathbf{F}, V\rangle \models Lp\ [y_{n0}]$ — and hence $\langle \mathbf{F}, V\rangle \models M(Lp \vee Lq)[x]$ — or, for each $n \in IN$, $\langle \mathbf{F}, V\rangle \models q\ [y_{n0}]$ or $\langle \mathbf{F}, V\rangle \models q\ [y_{n1}]$ — and hence $\langle \mathbf{F}, V\rangle \models Lq\ [z_f]$ for any f such that $\langle \mathbf{F}, V\rangle \models q\ [y_{nf(n)}]$ for all $n \in IN$; and, again, $\langle \mathbf{F}, V\rangle \models M(Lp \vee Lq)[x]$.

Next enters the familiar countable L_0-elementary subframe $\mathbf{F}'$ of $\mathbf{F}$ with a domain containing x and all y_{ni}'s. For any $z_f \in W - W'$, put $V(p) = \{y_{nf(n)} \mid n \in IN\}$ and $V(q) = W - V(p)$. Clearly, $\langle \mathbf{F}', V\rangle \models L(p \vee q)\ [x]$. Moreover, neither Lp nor Lq holds at any y_{ni}, and the same is true for any $z_g \in W'$. (Such a world z_g differs from z_f for at least one "argument" n; so Lp fails at z_g. But, it also yields the same value as z_f for at least one argument n, whence Lq fails at z_g as well.) In other words, $L(p \vee q) \rightarrow M(Lp \vee Lq)$ fails at x in $\mathbf{F}'$. ∨ QED.

To conclude this series of examples involving an application of the Löwenheim-Skolem theorem, a formula found in Gabbay [27], chapter 19, is treated by the same method. It is $\neg LM(\neg p \wedge Lp)$ or, equivalently, $ML(Lp \rightarrow p)$.

10.5 *Theorem.* $ML(Lp \rightarrow p) \notin M1$.

Proof. Consider the frame $\mathbf{F} = \langle W, R \rangle$ with $W = \{x\} \cup \{y_n, y_{ni} \mid n \in IN; i \in \{0, 1, 2\}\} \cup \{z_f, z_{fn} \mid n \in IN; f : IN \rightarrow \{0, 1\}\}$, and $R = \{\langle x, y_n \rangle, \langle y_n y_{n0} \rangle, \langle y_n, y_{n1} \rangle, \langle y_{n1}, y_{n2} \rangle \mid n \in IN\} \cup \{\langle x, z_f \rangle, \langle z_f, z_{fn} \rangle, \langle z_{fn}, y_{nf(n)} \rangle \mid n \in IN; f : IN \rightarrow \{0, 1\}\}$.



$\mathbf{F} \models ML(Lp \rightarrow p)[x]$. Suppose otherwise; say $\langle \mathbf{F}, V \rangle \models LM(Lp \wedge \neg p)[x]$ for some valuation V: a contradiction follows. Take $f : IN \rightarrow \{0, 1\}$ such that, for each $n \in IN$, $\langle \mathbf{F}, V \rangle \models \neg p\ [y_{nf(n)}]$. Then $\langle \mathbf{F}, V \rangle \models LM \neg p\ [z_f]$; which contradicts the fact that, by the assumption on x, $\langle \mathbf{F}, V \rangle \models M(Lp \wedge \neg p)\ [z_f]$.

If $\mathbf{F}'$ is a countable L_0-elementary subframe of $\mathbf{F}$ whose domain contains x and all worlds y_n and y_{ni}, then $ML(Lp \rightarrow p)$ fails at x in $\mathbf{F}'$. To see this, let $z_f \in W - W'$. (Note that no world z_{fn} will be in W'!) Set $V(p) = \{y_{n2} \mid n \in IN\} \cup \{y_{n(1-f(n))} \mid n \in IN\}$. An easy calculation shows that $\langle \mathbf{F}', V \rangle \models LM(Lp \wedge \neg p)[x]$. QED.

The methods of this chapter, together with those of the preceding one, would seem to provide the means for a characterization of $M1$. But, on the other hand, corollary 7.9 induced a strong suspicion that $M1$ is not syntactically characterizable at all. The most prudent course is, then, to consider special cases.

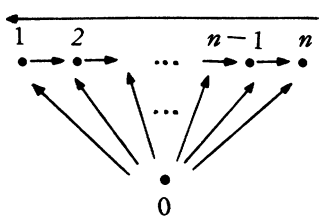
For a start, note that any modal formula is equivalent to one constructed using proposition letters, negations of proposition letters, $\wedge$, $\vee$, L and M. This point of view turned out to be advantageous with theorem 9.10 already. Now, let us restrict attention to the following subcases.

(1) (negations of) proposition letters, $\wedge$, $\vee$ and L. All formulas of this kind belong to $M1$, by theorem 9.10. (No combinations $M(\ldots L \ldots)$ or $M(\ldots\wedge\ldots)$ can occur.)

(2) (negations of) proposition letters, $\wedge$, $\vee$ and M. In this case, formulas outside of $M1$ are generated; such as $M(M\neg p \wedge \neg p) \vee M(Mp \wedge p)$, which is equivalent to $L(Lp \vee p) \rightarrow M(Mp \wedge p)$. (Cf. also theorem 10.3.)

10.6 *Lemma.* $L(Lp \vee p) \rightarrow M(Mp \wedge p) \notin M1$.

Proof. Consider the sequence of frames $\mathbf{F}_2$, $\mathbf{F}_3$, $\mathbf{F}_4, \ldots$, where $\mathbf{F}_n = \langle W_n, R_n \rangle$ with $W_n = \{0, 1, \ldots, n\}$ and $R_n = \{\langle 0, i\rangle \mid 1 \leqslant i \leqslant n.\cup \{\langle 1, 2\rangle, \ldots, \langle n-1, n\rangle, \langle n, 1\rangle\}$ $(n \geqslant 2)$.



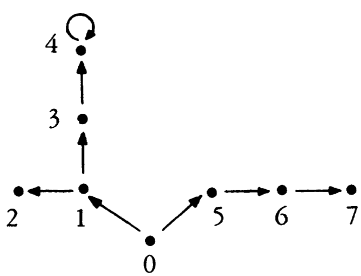
Note that $\mathbf{F}_n \models L(Lp \vee p) \rightarrow M(Mp \wedge p)[0]$ for all *odd* natural numbers n. Now suppose this formula were equivalent to some L_0-formula $\alpha(x)$. A routine argument, applying the compactness theorem to some suitably chosen set of L_0-sentences in combination with α, yields an infinite frame $\mathbf{F}$ with one world w without R-predecessors which is succeeded by infinitely many worlds, each having exactly one R-predecessor (apart from w) and exactly one R-successor. Furthermore, R is irreflexive and *no loops* of any finite length occur. But, then, it is easy to falsify $\alpha(x)$ at w in $\mathbf{F}$, by defining an "alternating" valuation V such that a world (other than w) belongs to $V(p)$ only if its R-successor and R-predecessor (other than w) do not belong to $V(p)$. QED.

To continue the above list:

(3) (negations of) proposition letters, $\vee$, L and M. This procedure still generates modal formulas outside of $M1$, such as $ML\neg p \vee MLp$; i.e., $LMp \rightarrow MLp$.

(4) (negations of) proposition letters, $\wedge$, L and M.

An answer for this case has not been found yet. It seems likely that all formulas generated in this way belong to $M1$; but the way to treat them is not clear. To get an impression, consider the modal formula $M(MLp \wedge MML\neg p)$. An obvious line of attack is by means of its closed substitution instances $M(ML\top \wedge MML\bot)$ (i.e., $M(M\top \wedge MML\bot)$, or $MMML\bot$) and $M(ML\bot \wedge MML\top)$ (i.e., $M(ML\bot \wedge MM\top))$. It is not implied by these, however; witness the following frame $\mathbf{F} = \langle W, R\rangle$.
$W = \{0, ..., 7\}$
$R = \{\langle 0, 1\rangle, \langle 1, 2\rangle, \langle 1, 3\rangle, \langle 3, 4\rangle, \langle 4, 4\rangle, \langle 0, 5\rangle, \langle 5, 6\rangle, \langle 6, 7\rangle\}$.



$MMML\bot$ and $M(ML\bot \wedge MM\top)$ hold at 0, but $M(MLp \wedge MML\neg p)$ does not: set $V(p) = \{4\}$. On the other hand, the latter formula does hold at 0 in $\langle W, R'\rangle$, with $R' = R \cup \{\langle 5, 4\rangle\}$. Is $M(MLp \wedge MML\neg p)$ in $M1$? Some trying out of similar examples has led to the conjecture that, at least, all formulas of this kind are *globally* equivalent to closed modal formulas.

A quite different type of restriction is the following.

10.7 *Definition.* A *modal reduction principle* is a modal formula of the form $\vec{M}p \rightarrow \vec{N}p$, where $\vec{M}$ and $\vec{N}$ are (possibly empty) sequences of modal operators L, M.

Many axioms used in modal logic are of this kind: they express interdependencies of iterated modal operators (cf. Chapter 5). The "negative" methods of this chapter and the "positive" method of substitutions of chapter 9 suffice for characterizing $M1$ in this special case. The next theorem, proven in [8], thus settles a problem of Fitch (cf.[26])- at least as far as local first-order definability is concerned.

10.8 *Theorem.* A modal reduction principle $\vec{M}p \rightarrow \vec{N}p$ is in $M1$ if and only if it has one of the following forms.

(i)	$M^iL^jp \rightarrow \vec{N}p$	for some $i, j \in IN$ and arbitrary $\vec{N}$
(ii)	$\vec{M}p \rightarrow L^iM^jp$	for some $i, j \in IN$ and arbitrary $\vec{M}$
(iii)	$L^i\vec{M}_1p \rightarrow \vec{N}_1\vec{M}_1p$	for some $i \in IN$ such that $\vec{N}_1$ has length i and for arbitrary $\vec{M}_1$
(iv)	$\vec{M}_1\vec{N}_1p \rightarrow M^i\vec{N}_1p$	for some $i \in IN$ such that M_1 has length i and for arbitrary $\vec{N}_1$.

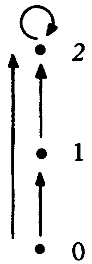
That modal formulas of these four kinds are in $M1$ is easy to see. Theorem 9.8 takes care of (i) and — after replacing p by $\neg p$ and contraposition — of (ii) as well; by corollary 9.2. Formulas of the forms (iii) and (iv) are equivalent to closed formulas and hence in $M1$; by lemma 9.6. (E.g., $LLMLp \rightarrow MLMLp$ is equivalent to $LLML\top \rightarrow MLML\top$, or even to $LLM\top \rightarrow M\top$.)

It remains to be seen, however, if modal reduction principles — or indeed modal formulas with only one proposition letter — are in any sense typical for the case of modal formulas in general. At least, the latter do *not* serve as a reduction class for the more general case. For, there are modal formulas with two proposition letters for which there is no modal formula with one proposition letter defining the same class of frames.

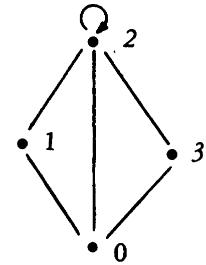
10.9 *Lemma.* $L((Lp \wedge p) \rightarrow q) \vee L(Lq \rightarrow p)$ is not equivalent to any modal formula with only one proposition letter.

Proof. Consider the two frames $\mathbf{F}_1$ and $\mathbf{F}_2$ with
$W_1 = \{0, 1, 2\}$, $R_1 = \{\langle 0, 1\rangle, \langle 0, 2\rangle, \langle 1, 2\rangle, \langle 2, 2\rangle\}$, and
$W_2 = \{0, 1, 2, 3\}$, $R_2 = \{\langle 0, 1\rangle, \langle 0, 2\rangle, \langle 0, 3\rangle, \langle 1, 2\rangle, \langle 3, 2\rangle, \langle 2, 2\rangle\}$.
As was shown in chapter 3, $E(L((Lp \wedge p) \rightarrow q) \vee L(Lq \rightarrow p), \forall y(Rxy \rightarrow \forall z(Rxz \rightarrow (Ryz \vee Rzy \vee z = y))))$; so this modal formula holds in $\mathbf{F}_1$, but not in $\mathbf{F}_2$. Now, it will be shown that, for any modal formula φ with

$\mathbf{F}_1$



$\mathbf{F}_2$



only one proposition letter (say p) such that $\mathbf{F}_1 \models \varphi$, $\mathbf{F}_2 \models \varphi$; from which fact the lemma follows.

Let V be a valuation on $\mathbf{F}_2$: it suffices to consider $V(p)$.

Case 1: 1 and 3 are both inside, or both outside $V(p)$. The function $f_1 = \{\langle 0, 0\rangle, \langle 3, 1\rangle, \langle 1, 1\rangle, \langle 2, 2\rangle\}$ is a *p-morphism* from $\langle \mathbf{F}_2, V\rangle$ onto $\langle \mathbf{F}_1, V'\rangle$ — where $V'(p) =_{def} V(p) - \{3\}$ — in the sense of definition 2.16. Therefore, any formula which is true at some world w in $\mathbf{F}_2$ is true at $f_1(w)$ in $\mathbf{F}_1$; by theorem 2.17.

Case 2: $1 \in V(p)$, $3 \notin V(p)$:

Subcase 2.1: $2 \in V(p)$. The function $f_2 = \{\langle 0, 0\rangle, \langle 3, 1\rangle, \langle 1, 2\rangle, \langle 2, 2\rangle\}$ is a p-morphism from $\langle \mathbf{F}_2, V\rangle$ onto $\langle \mathbf{F}_1, V'\rangle$ - where $V'(p) =_{def} V(p) - \{1\}$.

Subcase 2.2: $2 \notin V(p)$, is symmetrical: identify 3 and 2 instead of 1 and 2. The same holds for

Case 3: $1 \notin V(p)$, $3 \in V(p)$.

Finally, if $\mathbf{F}_2 \not\models \varphi$, then, for some valuation V on $\mathbf{F}_2$ and some $w \in W_2$, $\langle \mathbf{F}_2, V\rangle \models \neg\varphi[w]$. By the preceding paragraph, a valuation V' on $\mathbf{F}_1$ exists and a world $w' \in W_1$ such that $\langle \mathbf{F}_1, V'\rangle \models \neg\varphi[w']$. **QED.**

CHAPTER XI

RELATIVE FIRST-ORDER DEFINABILITY

In various applications of modal logic, it is natural to restrict attention to classes of frames satisfying some restriction on the alternative relation. E.g., in the original Kripke semantics for modal logic, *reflexivity* (i.e., $\forall x Rxx$) was required as a matter of course. Modulo such restrictions, the behaviour of modal formulas as regards first-order definability may change considerably. E.g., on the class of all reflexive frames, the formulas $L(Lp \vee p) \rightarrow M(Lp \wedge p)$, $ML(Lp \rightarrow p)$ and $L(Lp \vee p) \rightarrow M(Mp \wedge p)$ of the preceding chapter would become universally valid; and hence trivially first-order definable. (The first example may not be obvious. Note that $L\varphi \rightarrow \varphi$ and $\varphi \rightarrow M\varphi$ have become universally valid. Then $L(Lp \vee p) \rightarrow L(p \vee p) \rightarrow Lp \rightarrow Lp \wedge p \rightarrow M(Lp \wedge p)$.) Other restrictions which have been considered are *succession* (i.e., $\forall x \exists y Rxy$) and *transitivity* (i.e., $\forall x \forall y (Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))$). The second of these will be considered in more detail later on. As for the first, the following example, taken from [8], shows how such a simple restriction may affect previous results like theorem 10.8.

11.1 *Theorem.* On the class of frames satisfying $\forall x \exists y Rxy$, a modal reduction principle is in $M1$ if and only if it has one of the following forms.

(i) $M^i L^j p \rightarrow \vec{N}p$,
(ii) $\vec{M}p \rightarrow L^i M^j p$, or
(iii) $\vec{M}p \rightarrow \vec{N}p$, where $\vec{M}$ and $\vec{N}$ are equally long sequences of modal operators such that, for each $k \in IN$, if $(\vec{M})_k = M$, then $(\vec{N})_k = M$.

It is of some interest to determine not too restrictive conditions under which all modal formulas become first-order definable. To this end, recall lemma 9.7: modal formulas of degree at most one belong to $M1$. Thus, one way to make all modal formulas first-order definable is by introducing relational conditions which reduce every modal formula to one of degree at most one. This means that, amongst others, sequences of modal operators will have to collapse to single such operators. One obvious principle is, therefore, $LLp \leftrightarrow Lp$ (and hence $MMp \leftrightarrow Mp$). Moreover, alternations will have to be treated; such as "ML". In view of existing modal logics like $S5$ (cf. chapter 5), $MLp \leftrightarrow Lp$ (and hence also $LMp \leftrightarrow Mp$) seems to be an interesting candidate. Another type of nesting of modal operators occurs in, e.g., $L(p \vee Lq)$. How are these to be contracted? Fortunately, no special provision is needed for this case. As was noted in [8], the formula $MLp \rightarrow LLp$ — which is derivable from the above two — implies the general reduction principle

$$L(p \vee Lq \vee Mr) \leftrightarrow (Lp \vee LLq \vee LMr).$$

(A sketch of the proof goes as follows:
$L(p \vee Lq \vee Mr) \rightarrow (MLq \vee L(p \vee Mr)) \rightarrow (LLq \vee L(p \vee Mr)) \rightarrow (LLq \vee (Lp \vee MMr)) \rightarrow (LLq \vee Lp \vee LMr)$.)
From these facts, the following result may be derived.

11.2 *Lemma.* For any modal formula φ, there exists a modal formula ψ of degree at most 1, which can be constructed effectively from φ, such that $\varphi \leftrightarrow \psi$ is universally valid on the class of all frames in which both $LLp \leftrightarrow Lp$ and $MLp \leftrightarrow Lp$ hold.

Using the proof of theorem 9.8, it is easy to discover that

$$\begin{array}{ll} E(LLp \rightarrow Lp, & \forall y(Rxy \rightarrow \exists z(Rxz \wedge Rzy))) \\ E(Lp \rightarrow LLp, & \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))) \\ E(MLp \rightarrow Lp, & \forall y(Rxy \rightarrow \forall z(Rxz \rightarrow Ryz))) \\ E(Lp \rightarrow MLp, & \exists y(Rxy \wedge \forall z(Ryz \rightarrow Rxz))). \end{array}$$

Some reflection shows that the first condition follows from the third, and that the fourth may be replaced by $\exists yRxy$, given the second. In other words, the following has been proven.

11.3 *Theorem.* On the class of frames satisfying $\forall x \exists yRxy$,

$\forall x \forall y(Rxy \to \forall z(Ryz \to Rxz))$ and $\forall x\ \forall y(Rxy \to \forall z(Rxz \to Ryz))$, any modal formula belongs to *M1*.

This result implies that each modal formula is first-order definable on the basis of *S5* (which requires R to be an *equivalence relation*); but it is slightly stronger in that certain frames whose alternative relation is not an equivalence relation satisfy the present condition. The frame $\langle\{0, 1\}, \{\langle 0, 1\rangle, \langle 1, 1\rangle\}\rangle$ provides an example.

Probably the most interesting condition on frames is that the alternative relation be *transitive*. Because the class of frames satisfying this condition is defined by a modal formula, viz. $Lp \to LLp$, it is, conveniently, closed under the basic operations of taking generated subframes, disjoint unions and *p*-morphic images. Moreover, its complement is closed under ultrafilter extensions. (Cf. the relevant theorems in chapter 2.) As $Lp \to LLp \in M1$, the class of transitive frames is even closed under ultrafilter extensions itself; by theorem 8.9. In other words, the model theory of preceding chapters is applicable without reservations.

The difference between *M1* and $\bar{M}1$ virtually disappears modulo transitivity, as will be seen in theorem 11.8 below. First, a few auxiliary results are needed. It will be proven in chapter 15 that, if φ is any modal formula in $\bar{M}1$, then there exists a *restricted* L_0-formula $\alpha = \alpha(x)$ (i.e., all quantifiers in α occur only in contexts $\forall y(Rzy \to$ or $\exists y(Rzy \wedge$; where z is distinct from y) such that $\bar{E}(\varphi, \forall x\ \alpha)$. For the special case of transitive frames, this result is easy to prove (cf. corollary 11.7). The utility of restricted L_0-formulas derives from the following property (cf. also the proof of lemma 3.11).

11.4 *Lemma.* If $\alpha = \alpha(x_1, \ldots, x_n)$ is a restricted L_0-formula, and if $\mathbf{F}_1$ is a generated subframe of $\mathbf{F}_2$ with $w_1, \ldots, w_n \in W_1$, then $\mathbf{F}_1 \models \alpha[w_1, \ldots, w_n]$ if and only if $\mathbf{F}_2 \models \alpha[w_1, \ldots, w_n]$.

This *invariance* of restricted formulas for *generated subframes* follows by a routine formula induction. Note that, in particular, $\mathbf{F} \models \alpha[w, w_1, \ldots, w_n]$ if and only if $TC(\mathbf{F}, w) \models \alpha[w, w_1, \ldots, w_n]$ for restricted formulas $\alpha = \alpha[x, x_1, \ldots, x_n]$.

11.5 *Definition.* Let x be some fixed variable. For any L_0-formula α,

$R_x(\alpha)$ is defined inductively according to the clauses

$$\begin{aligned} R_x(\alpha) &= \alpha \text{ for atomic formulas } \alpha \\ R_x(\neg\alpha) &= \neg R_x(\alpha) \\ R_x(\alpha \to \beta) &= R_x(\alpha) \to R_x(\beta) \\ R_x(\forall y\, \alpha) &= R_x([x/y]\ \alpha) \wedge \forall y(Rxy \to R_x(\alpha)), \end{aligned}$$

where $[x/y]\ \alpha$ is the result of substituting x for y in α, taking bound alphabetic variants when needed, and where y is distinct from x (or made to be distinct from x, if necessary).

Clearly, all formulas of the form $R_x(\alpha)$ are restricted. Moreover, the following connection exists between α and $R_x(\alpha)$.

11.6 *Lemma.* If F is a transitive frame $\langle W, R\rangle$ with $w, w_1, ..., w_n \in TC(\mathsf{F}, w)$, then, for any L_0-formula $\alpha = \alpha(x, x_1, ..., x_n)$, $\mathsf{F} \models R_x(\alpha)[w, w_1, ..., w_n]$ if and only if $TC(\mathsf{F}, w) \models \alpha[w, w_1, ..., w_n]$.

The proof is by induction on α, noting that the domain of $TC(\mathsf{F}, w)$ is $\{w\} \cup \{v \in W \mid Rwv\}$.

11.7 *Corollary.* If $\bar{E}(\varphi, \alpha)$, then $\bar{E}(\varphi, \forall x R_x(\alpha))$.

Proof. If $\mathsf{F} \models \varphi$, then — by corollary 2.12 — for all $w \in W$, $TC(\mathsf{F}, w) \models \varphi$; whence, for all $w \in W$, $TC(\mathsf{F}, w) \models \alpha$. Therefore, by lemma 11.6, $\mathsf{F} \models \forall x R_x(\alpha)$.

If $\mathsf{F} \models \forall x R_x(\alpha)$, then, for all $w \in W$, $\mathsf{F} \models R_x(\alpha)[w]$ (and hence $TC(\mathsf{F}, w) \models \alpha[w]$ and so $TC(\mathsf{F}, w) \models \alpha$; because α is an L_0-*sentence*). It follows that $TC(\mathsf{F},w) \models \varphi$ for all $w \in W$, and, therefore, that $\mathsf{F} \models \varphi$; again by corollary 2.12. **QED**.

Now comes the main result.

11.8 *Theorem.* On the transitive frames, for any modal formula φ, $\varphi \in \bar{M}1$ if and only if $L\varphi \in M1$.

Proof. First, suppose that $\varphi \in \bar{M}1$; say $\bar{E}(\varphi, \alpha)$. By 11.7, it may be assumed that α is of the form $\forall x\ \beta$, where $\beta = \beta(x)$ is a restricted L_0-formula. Let y be any variable not occurring in $\forall x\ \beta$. It will be shown that

(*) $\quad E(L\varphi, \forall y(Rxy \to [y/x]\ \beta))$.

To see this, let $\mathbf{F} \models L\,\varphi[w]$. It follows that $TC(\mathbf{F}, w) \models L\,\varphi[w]$; or, for all $v \in W$ such that Rwv, $TC(\mathbf{F}, w) \models \varphi[v]$. Call the subframe of $TC(\mathbf{F}, w)$ obtained by restricting the domain to $\{v \in W \mid Rwv\}$: $TC(\mathbf{F}, w)^{-}$. (w may have been dropped in this way.) Note that $TC(\mathbf{F}, w)^{-}$ is a generated subframe of $TC(\mathbf{F}, w)$. Therefore, $TC(\mathbf{F}, w)^{-} \models \varphi[v]$ for each v in its domain; and hence $TC(\mathbf{F}, w)^{-} \models \varphi$. But, then, $TC(\mathbf{F}, w)^{-} \models \forall x\, \beta$; and $TC(\mathbf{F}, w)^{-} \models \beta[v]$ for all R-successors v of w. Now, $TC(\mathbf{F}, w)^{-}$, being a generated subframe of $TC(\mathbf{F}, w)$, is also a generated subframe of $\mathbf{F}$; and, therefore —using lemma 11.4 — $\mathbf{F} \models' \beta[v]$ for all R-successors v of w. In other words, $\mathbf{F} \models \forall y(Rxy \rightarrow [y/x]\ \beta)[w]$. This establishes one direction of (*). The other direction follows by inverting this argument. Thus $L\varphi \in M1$.

Next, suppose that $L\varphi \in M1$. Then, obviously, $L\varphi \in \bar{M}1$. An argument analogous to the one proving lemma 8.8 shows that, then, $\varphi \in \bar{M}1$. In fact, this implication holds generally. Suppose that $\varphi \notin \bar{M}1$. Then, by theorem 8.6, there exists a set of frames $\{\mathbf{F}_i \mid i \in I\}$ and an ultrafilter U on I such that $\mathbf{F}_i \models \varphi$ for all $i \in I$, but $\Pi_U \mathbf{F}_i \not\models \varphi$. Define frames $\mathbf{F}_i'$ by adding some new world a to W_i and setting $\bar{R}_i = R_i \cup \{\langle a, w\rangle \mid w \in W_i\}$. Clearly, if $\mathbf{F}_i \models \varphi$, then $\mathbf{F}_i' \models L\varphi$. Moreover, $\Pi_U \mathbf{F}_i$ is a generated subframe of $\Pi_U \mathbf{F}_i'$ in which φ fails at some w. Since in $\Pi_U \mathbf{F}_i'$ $R((\langle a\rangle_{i \in I})_U\, w$ holds (use Łoś' Theorem), $L\varphi$ fails at $(\langle a\rangle_{i\in I})_U$. Therefore, $\Pi_U \mathbf{F}_i' \not\models L\varphi$; and hence $L\varphi \notin \bar{M}1$, $L\varphi$ not being preserved under ultrapowers. QED.

It is an open question if actually $M1 = \bar{M}1$ on the transitive frames.

Transitivity allows for various reductions in complexity. E.g., $LMLMp \leftrightarrow LMp$ and $LMMp \leftrightarrow LMp$ are universally valid in this case. This leaves only sequences of modal operators of the forms

M^i, M^iL, M^iLM, L^i, L^iM and L^iML.

(Use these two reduction principles, as well as their duals $MLMLp \leftrightarrow MLp$ and $MLLp \leftrightarrow MLp$. Cf. [8].) For modal reduction principles, this just leaves the types

(1)	M^iLM, M^jL	$(j \geqslant 1)$
(2)	M^iLM, M^jLM	$(j \geqslant 1)$
(3)	M^iLM, L^jML	
(4)	L^iM, M^jL	$(i \geqslant 1,\ j \geqslant 1)$
(5)	L^iM, M^jLM	$(i \geqslant 1,\ j \geqslant 1)$
	L^iM, L^jML	= type (1), by an obvious conversion,

$L^iML,\ M^jL$ = type (5), by an obvious conversion,

(6) $L^iML,\ M^jLM$
$L^iML,\ L^jML$ = type (2), by an obvious conversion.

These are the only potential candidates for non-first-order definability: all other types are first-order definable, by theorem 10.8. But, in fact all modal reduction principles of these six types are first-order definable on transitive frames; with equivalents as given in the following table (cf. [8]).

(1) $\forall y((R^ixy \wedge \forall z(Ryz \rightarrow \exists uRzu)) \rightarrow \exists v(R^jxv \wedge \forall w(Rvw \rightarrow (v = w \wedge Ryv))))$. (Cf. theorem 7.4 for the special case of $LMp \rightarrow MLp$.)
(2) If $i \geqslant j$: $x = x$. If $i < j$, then
$\forall y(R^ixy \rightarrow \exists zRyz) \vee \exists y(R^jxy \wedge \neg \exists zRyz)$.
(3) $\forall y((R^ixy \wedge \forall z(Ryz \rightarrow \exists uRzu)) \rightarrow \forall v(R^jxv \rightarrow \exists w(Rvw \wedge \forall s(Rws \rightarrow (s = w \wedge Ryw)))))$.
(4) $\exists y(R^ixy \wedge \neg \exists zRyz) \vee \exists y(R^jxy \wedge \forall z(Ryz \rightarrow z = y))$.
(5) $\exists y(R^ixy \wedge \neg \exists zRyz) \vee \exists y(R^jxy \wedge \forall z(Ryz \rightarrow \exists uRzu))$.
(6) $\exists y(R^ixy \wedge \neg \exists zRyz) \vee \exists y(R^jxy \wedge \forall z(Ryz \rightarrow \exists uRzu))$.

This exhausts all possibilities:

11.9 *Theorem.* On the transitive frames, all modal reduction principles are in *M1*.

It does not follow that all modal formulas are in $\bar{M}1$ modulo transitivity. E.g., $L(Lp \rightarrow p) \rightarrow Lp$ ("Löb's Formula") still defines the class of frames whose converse relation is well-founded; which is not elementary. This example was presented in chapter 3, as was another one, viz. $L(L(p \rightarrow Lp) \rightarrow p) \rightarrow p$ ("Dummett's Formula"), which is not even first-order definable on the class of frames which are *transitive*, *reflexive* and *connected*. The latter formula is of interest because of its connection with intuitionistic logic. It axiomatizes the strongest modal logic (stronger than, e.g., *S4*!) for which Gödel's embedding of intuitionistic logic into modal logic works. (In connection with chapter 6, it should be noted that Dummett's Formula is complete with respect to the class of finite reflexive trees; whereas Löb's Formula is complete with respect to the class of finite irreflexive trees. Cf. Segerberg [67].)

A quite different reduction in complexity is obtained as follows. Let φ be any modal formula. List its subformulas as $\varphi_1, \ldots, \varphi_k$ and take new proposition letters $q_1, \ldots, q_m$ not occurring in φ for those formulas φ_i which are not proposition letters themselves. Let $Q_i = q_i$ if φ_i was not a proposition letter; $Q_i = \varphi_i$, otherwise. Then form the conjunction of all "connecting formulas". I.e.,

if $\varphi_i = \neg\varphi_j$, then take $(Q_i \leftrightarrow \neg Q_j) \wedge L(Q_i \leftrightarrow \neg Q_j)$,
if $\varphi_i = \varphi_j \rightarrow \varphi_k$, then take $(Q_i \leftrightarrow (Q_j \rightarrow Q_k)) \wedge L(Q_i \leftrightarrow (Q_j \rightarrow Q_k))$, and
if $\varphi_i = L\varphi_j$, then take $(Q_i \leftrightarrow LQ_j) \wedge L(Q_i \leftrightarrow LQ_j)$.
Call this conjunction $Q(\varphi)$.

E.g., $Q(LLLp \rightarrow p) = (q_1 \leftrightarrow Lp) \wedge L(q_1 \leftrightarrow Lp) \wedge (q_2 \leftrightarrow Lq_1) \wedge L(q_2 \leftrightarrow Lq_1) \wedge (q_3 \leftrightarrow Lq_2) \wedge L(q_3 \leftrightarrow Lq_2) \wedge (q_4 \leftrightarrow (q_3 \rightarrow p)) \wedge L(q_4 \leftrightarrow (q_3 \rightarrow p))$. Note that the degree of $Q(\varphi)$ is at most two! The following result is easy to establish, once one has understood the trick behind the definition of $Q(\varphi)$.

11.10 *Lemma.* For all transitive general frames $\langle \mathsf{F}, W \rangle$ and all $w \in W$, and for all modal formulas φ,

$$\langle \mathsf{F}, W \rangle \models \varphi[w] \text{ if and only if } \langle \mathsf{F}, W \rangle \models Q(\varphi) \rightarrow q[w],$$

where q is the proposition letter in the above construction which corresponds to φ itself.

11.11 *Corollary.* For any modal formula φ, a modal formula ψ of degree at most two can be constructed effectively from φ, such that, on the class of transitive frames,

$$\varphi \in M1 \text{ if and only if } \psi \in M1.$$

In other words, it suffices to consider formulas of degree 2 for the study of first-order definability on the transitive frames. (Formulas of degree 1 are first-order definable anyway.) But, in fact, the above idea may also be used to show that it suffices to consider modal logics with axioms of degree at most 2 in the Completeness Theory on transitive frames.

The best-known example of a relational condition on frames occurs in *intuitionistic logic.* There,one restricts attention to reflexive and transitive frames. (Some authors add *anti-symmetry; i.e.,* $\forall x \forall y((Rxy \wedge Ryx) \rightarrow x = y)$.) But, there is another restriction as well. Valuations V have to satisfy the requirement of closure under R-successors for their values $V(p)$. This makes intuitionistic formulas better-behaved than modal ones in several ways. E.g., (cf. Smorynski [73]):

(i) For all frames $\mathbf{F} = \langle W, R\rangle$, all $w \in W$ and all valuations V on $\mathbf{F}$, if $\langle \mathbf{F}, V\rangle \models \varphi[w]$, then $\langle \mathbf{F}, V\rangle \models \varphi[v]$ for all R-successors v of w. (In other words, $\varphi \to L\varphi$ becomes universally valid.)

(ii) For all frames $\mathbf{F} = \langle W,R\rangle$, all $w \in W$ and all valuations V on $\mathbf{F}$, if $\langle \mathbf{F}, V\rangle \models \varphi[w]$, then a *finite* submodel $\mathbf{M}$ of $\langle \mathbf{F}, V\rangle$ exists such that $\mathbf{M} \models \varphi[w]$.

(i) does not hold for modal formulas in general. (Consider e.g., $\neg p$.) Inspection of Smorynski's proof shows that (ii) does hold for all modal formulas, given this kind of frame and valuation. (It does not hold for arbitrary valuations, however. E.g., if $\mathbf{F} = \langle IN, \leqslant\rangle$ and $V(p)$ is the set of odd numbers, then $\langle \mathbf{F}, V\rangle \models LMp \wedge LM\neg p[0]$; but this formula holds at 0 in no finite submodel of $\langle \mathbf{F}, V\rangle$.)

The first question to be answered in this area is:
"Are there intuitionistic formulas without first-order equivalents?"

We suspect the answer to be negative. (No doubt this conjecture must have been formulated already by persons working in the field of intermediate logics.) To get an impression of the kind of formula being discussed here, consider the intuitionistic principles

(1) $(p \to q) \vee (q \to p)$ (i.e., modally, $L(Lp \to Lq) \vee L(Lq \to Lp)$)
defined by $\forall y(Rxy \to \forall z(Rxz \to (Ryz \vee Rzy)))$.
(2) $(\neg p \to (q \vee r)) \to ((\neg p \to q) \vee (\neg p \to r))$:
"Kreisel & Putnam's Formula"; i.e., modally,
$L(L(L\neg Lp \to (Lq \vee Lr)) \to (L(L\neg Lp \to Lq) \vee L(L\neg Lp \to Lr)))$,
defined by
$\neg\ \exists y\ \exists z(Rxy \wedge Rxz \wedge \neg Ryz \wedge \neg Rzy \wedge \forall u((Rxu \wedge Ruy \wedge Ruz) \to \exists\ v(Ruv \wedge \neg Ryv \wedge \neg Rzv)))$.
(3) $((\neg\neg p \to p) \to (p \vee \neg p)) \to (\neg\neg p \vee \neg p)$: "Scott's Formula", for which no first-order equivalent has been found yet.

In the meantime, the above question has been answered in the affirmative: Scott's Formula is not first-order definable. But, the conjecture is true for all *disjunction-free* formulas. More systematic information is found in P. Rodenburg, *Intuitionistic Correspondence Theory*, Mathematisch Instituut, Universiteit van Amsterdam, 1982.

Another possible field for relative Correspondence Theory is the semantics of *relevance logics*, in terms of a *ternary* relation between worlds, satisfying certain restrictions (and with valuations subject to a condition similar to the above). Cf. [27], or the more general type of question in [38]. Further examples, such as *conditional logic* or *quantum logic*, are given in the "Correspondence Theory" chapter of the *Handbook of Philosophical Logic* (D. Gabbay & F. Guenthner, eds; Reidel, Dordrecht, 1984, volume II).

CHAPTER XII

MODAL PREDICATE LOGIC

This chapter forms a kind of appendix to the main theme of this book. It serves to point out how a generalization of Correspondence Theory to the richer language of modal predicate logic (instead of just modal *propositional* logic) could be effected.

The first choice to be made is that of a suitable modal predicate language. For simplicity, attention is restricted here to an ordinary predicate-logical language with individual variables and predicate constants, without other kinds of variables or constants. Identity is not included. As is well-known, cf. Hughes & Cresswell [36], individual constants and identity cause (philosophical) troubles in this area. To this predicate-logical language, the modal operators L and M are added as unary operators from formulas to formulas. The main novelty which arises is the interplay of individual quantifiers and modal operators, as in the famous *Barcan Formula* $\forall xLAx \rightarrow L\forall xAx$.

On the semantic side, worlds will now get some inner structure: they have "domains of individuals" on which the predicates are interpreted. No universally accepted semantics exists; and the one to be given here is only intended as a basis for technical discussion, not as a philosophical paradigm.

12.1 *Definition.* A *frame* is an ordered couple $\langle W, R\rangle$ with a non-empty set W (of "worlds") and a binary relation R on W ("accessibility"). A *skeleton* is an ordered triple $\mathbf{S} = \langle W, R, D\rangle$, where $\langle W, R\rangle$ is a frame and D a function with domain W assigning, to each world $w \in W$, a non-empty set D_w ("w's domain of individuals"). A *model* is an ordered quadruple $\langle W, R, D, I\rangle$, where $\langle W, R, D\rangle$ is a skeleton and I is a function with domain W assigning, to each world $w \in W$, an interpretation I_w of the predicate-logical language on the domain D_w.

Frames give the abstract pattern of the set of worlds. Skeletons contain information about the presence of individuals as well: they will be the basic structures of the new Correspondence Theory. Note that different worlds may have the same domain of individuals. Defining skeletons as couples $\langle D, R\rangle$, letting R relate non-empty sets of individuals, is therefore less suitable, since it would mean losing some power of discrimination. Finally, models add interpretations to make formulas of modal predicate logic evaluable.

12.2 *Definition.* Let φ be a formula of modal predicate logic. Let $\mathbf{M} = \langle W, R, D, I\rangle$ be a model. $\mathbf{M} \models \varphi[w, A]$ ("*φ holds in* $\mathbf{M}$ *at w under A*") is defined inductively, for all worlds $w \in W$ and assignments A mapping x_i onto $A(x_i) \in D_w$ $(1 \leqslant i \leqslant n)$, according to the clauses:

(i) $\mathbf{M} \models Py_1 \ldots y_m[w, A]$ iff $\langle A(y_1), \ldots, A(y_m)\rangle \in I_w(P)$, for any m-ary predicate constant P and m-tuple of individual variables $y_1, \ldots, y_m$ $(m \geqslant 1)$

(ii) $\mathbf{M} \models \neg\varphi[w, A]$ iff *not* $\mathbf{M} \models \varphi[w, A]$

(iii) $\mathbf{M} \models \varphi \rightarrow \psi\, [w, A]$ iff *if* $\mathbf{M} \models \varphi[w, A]$, *then* $\mathbf{M} \models \psi[w, A]$

(iv) $\mathbf{M} \models \forall x\varphi[w, A]$ iff *for all* assignments A' different from A at most in that $A'(x) \neq A(x)$, where $A'(x)$ is to be *in* D_w, $\mathbf{M} \models \varphi[w, A']$

(v) $\mathbf{M} \models L\varphi[w, A]$ iff *for all* $v \in W$ such that Rwv *and* $A(y) \in D_v$ for all free variables y of φ, $\mathbf{M} \models \varphi[v, A]$.

Note that many alternative clauses are possible. Here, a specific choice has been made which, according to many authors, has a certain naturality. Notably, the universal quantifier of (iv) ranges over all objects *in the given world*; i.e., in philosophical jargon, quantification is over *existents*, not *subsistents*. Another way to implement this technically would be by having one single domain D, common to all worlds in W, and enriching the language with an *existence* predicate E true at w of exactly those $d \in D$ which "exist" at w.

Then, the necessity operator of (v) ranges over all R-successors of a given world w which contain the objects $A(y)$ with respect to which φ is being evaluated. (To use an ordinary language example, "She is always angry" does not mean that, at all points in time, she is angry; but that she is angry at all points in time *during her existence*.) Again, an existence predicate would provide an alternative way of obtaining this interpretation, even within a language with unrestricted necessity (L). Finally, a dual clause (dual to (v)) for M is easily formulated:

(vi) $\mathbf{M} \models M\varphi[w, A]$ iff *there exists a* $v \in W$ such that Rwv *and* $A(y) \in D_v$ for all free variables y of φ and $\mathbf{M} \models \varphi[v, A]$.

A convenient predicate-logical language for describing this semantics has two *sorts*, one for individuals and one for worlds. There are a few distinguished predicates; viz. R (denoting a binary relation between objects of the second sort) and E (denoting a binary relation between objects of the first sort and objects of the second sort). Later on, identity ($=$) between objects of the second sort will be useful as well. Moreover, the language has, for any m-ary predicate constant P of the original language of modal predicate logic, an $m + 1$-ary predicate constant P^* denoting the predicate P relativized to worlds. Much as in chapter 3, a translation may now be defined taking modal formulas to formulas of this two-sorted predicate-logical language.

12.3 *Definition* Let u be some fixed variable of the second sort.

(i) $ST(Py_1 \ldots y_m) = P^*uy_1 \ldots y_m$
(ii) $ST(\neg\varphi) = \neg ST(\varphi)$
(iii) $ST(\varphi \rightarrow \psi) = ST(\varphi) \rightarrow ST(\psi)$
(iv) $ST(\forall x\varphi) = \forall x(Exu \rightarrow ST(\varphi))$
(v) $ST(L\varphi) = \forall v(Ruv \rightarrow ((Ey_1v \wedge \ldots \wedge Ey_mv) \rightarrow [v/u]\, ST(\varphi))$,

where v is a variable over objects of the second sort which is distinct from u and does not occur in $ST(\varphi)$, while $y_1, \ldots, y_m$ are the free individual variables of φ.

The conclusions which may be drawn from the existence (and properties) of this translation may be supplied by the reader who has understood chapter 3. Correspondence Theory proper arises when the following notions are introduced.

12.4 *Definition.* For a skeleton $\mathbf{S} = \langle W, R, D\rangle$ and $w \in W$, $\mathbf{S} \models \varphi[w]$

("φ *holds in* $\mathbf{S}$ *at* w") if, for all models $\mathbf{M} = \langle W, R, D, I\rangle$, $\mathbf{M} \models \varphi[w]$.

In this definition, φ is supposed to be a *sentence* of modal predicate logic; which allows us to abstract from assignments A in the formulation. Note that skeletons are structures for the two-sorted predicate logic L_0^+ whose only predicates are R, $=$ and E as introduced above.

12.5 *Definition.* For a sentence φ of modal predicate logic, and for a formula α of L_0^+ with one free, world variable, $E^+(\varphi, \alpha)$ holds if, for all skeletons $\mathbf{S} \models \langle W, R, D\rangle$ and all $w \in W$, $\mathbf{S} \models \varphi[w]$ if and only if $\mathbf{S} \models \alpha[w]$. $M1^+$ is the set of all sentences φ for which a formula α exists such that $E^+(\varphi, \alpha)$.

An example of this notion is provided by the Barcan Formula:

$$E(\forall xLAx \rightarrow L\forall xAx,\ \forall v(Ruv \rightarrow \forall y(Eyv \rightarrow Eyu))),$$

as is easy to check. Before giving more examples, we formulate a *conservation result*. Recall definition 3.11.

12.6 *Theorem.* There exists an effective translation from sentences φ in modal predicate logic to modal propositional formulas $mp(\varphi)$ such that, if $E^+(\varphi, \alpha)$, where α does not contain E, then $E(mp(\varphi), \alpha)$.

Proof. Define mp as simply "stripping away" predicate-logical material. Formally, let $mp(Py_1 \ldots y_m) = P$ (with P now regarded as a proposition letter), $mp(\neg\varphi) = \neg mp(\varphi)$, $mp(\varphi \rightarrow \psi) = mp(\varphi) \rightarrow mp(\psi)$, $mp(\forall x\varphi) = mp(\varphi)$ and $mp(L\varphi) = Lmp(\varphi)$. Now, suppose that $E^+(\varphi, \alpha)$, where α is as described. It is to be shown that $E(mp(\varphi), \alpha)$.

First, let $\mathbf{F}$ $(= \langle W, R\rangle) \models \alpha[w]$. By assumption, $\langle W, R, D\rangle \models \varphi[w]$ will hold; regardless of the choice of D. Choose D', then, as a function assigning the same singleton (say) $\{a\}$ to each $w \in W$. It was to be shown that $\mathbf{F} \models mp(\varphi)[w]$; i.e., for any valuation V on $\mathbf{F}$, $\langle \mathbf{F}, V\rangle \models mp(\varphi)[w]$. So, let V be an arbitrary valuation on $\mathbf{F}$. Define an interpretation I as follows. I_w assigns, to an m-ary predicate constant P, $I_w(P) = \{\langle a, \ldots (m \text{ times}) \ldots, a\rangle\}$ if $w \in V(p)$ for the corresponding proposition letter p; $I_w(P) = \emptyset$, otherwise. Then, since $\langle W, R, D', I\rangle \models \varphi[w]$, it also holds that $\langle \mathbf{F}, V\rangle \models mp(\varphi)[w]$.

Next, let $\mathbf{F}$ $(= \langle W, R\rangle) \models mp(\varphi)[w]$. It is to be shown that $\mathbf{F} \models$

$\alpha[w]$. It suffices to show that $\langle W, R, D'\rangle \models \varphi[w]$ for the above D'; because of $E^+(\varphi, \alpha)$. Let, then, I be any interpretation turning this skeleton into the model $\langle W, R, D', I\rangle$. Define $V(p)$ as the set of those worlds $w \in W$ for which $I_w(P)$ is non-empty. Obviously, for this valuation, $\langle \mathbf{F}, V\rangle \models mp(\varphi)[w]$; and it follows as before that $\langle W, R, D', I\rangle \models \varphi[w]$. **QED.**

In other words, one need not expect new L_0-formulas to become modally definable.

Probably the most important modal principles in this area (apart from the Barcan Formula) are the interchange principles

(1) $L\forall xAx \to \forall xLAx$
(2) $\exists xLAx \to L\exists xAx$
(3) $L\exists xAx \to \exists xLAx$.

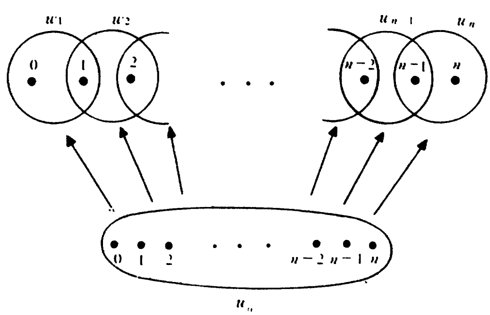
For the first two of these, the following equivalences may be proven:

$E^+(L\forall xAx \to \forall xLAx, u = u)$
$E^+(\exists xLAx \to L\exists xAx, \forall y(Eyu \to \forall v(Ruv \to Eyv)))$.

The third is more complicated.

12.7 *Lemma.* $L\exists xAx \to \exists xLAx \notin M1^+$.

Proof. Suppose that $E^+(L\exists xAx \to \exists xLAx, \alpha)$ for some L_0^+-formula α. Using the compactness theorem, a contradiction may be derived from this assumption. To see this, consider skeletons $\mathbf{S}_n = \langle W_n, R_n, D_n\rangle$ with $W_n = \{w_0, ..., w_n\}$, $R_n = \{\langle w_0, w_i\rangle \mid 1 \leqslant i \leqslant n\}$ and D_n given by $D_n(w_0) = \{0, ..., n\}$, $D_n(w_i) = \{i-1, i\}$ $(1 \leqslant i \leqslant n)$.



Note that $\mathbf{S}_n \models L\exists xAx \rightarrow \exists xLAx[w_0]$. (Intuitively, if $L\exists xAx$ holds in w_0, then xAx holds in w_1. Now, in w_1, either Ax holds of 0 — and hence LAx holds of 0 in w_0 — or Ax holds of 1. Then, either Ax holds of 1 in w_2 as well — in which case LAx holds of 1 in w_0 — or Ax holds of 2 in w_2. Etc. Continuing this argument, it is found that either LAx has become true in w_0 for some $k < n$; or Ax has to hold of n in w_n, and $\exists xLAx$ is true in w_0 after all (because LAx holds of n).) On the other hand, if the sequence $w_1, w_2, \ldots, w_n, \ldots$ were to continue indefinitely towards the right, then $L\exists xAx \rightarrow \exists xLAx$ could be falsified at w_0 in an obvious manner. Moreover, if the chain of R-successors of u were to extend indefinitely in both directions, it would be falsifiable as well.

These observations yield the required argument; an outline of which follows here. Let the purported L_0^+-equivalent α of $L\exists xAx \rightarrow \exists xLAx$ have the free world variable u. Consider the set Σ of L_0^+-formulas consisting of α together with formulas expressing that

(i) all R-successors of u have a two-element domain,
(ii) all domains of R-successors of u are contained in the domain of u,
(iii) if the domains of two distinct R-successors of u have an object in common, they have only one object in common,
(iv) there are exactly two R-successors of u which have an overlapping domain with exactly one other R-successor of u; all others have an overlapping domain with exactly two other R-successors of u,
$(\text{v})_n$ there are at least n R-successors of u $(n \geqslant 2)$.

Every finite subset of Σ is satisfiable, but Σ itself is not satisfiable: which is the required contradiction. For, in any model of Σ, u will be succeeded by a set of R-successors whose domains are such that they allow for a counter-example to $L\exists xAx \rightarrow \exists xLAx$, given in the way indicated above. QED.

Lemma 12.7 should be compared with lemma 9.7 and theorem 10.2. In the field of modal *predicate* logic, formulas of modal degree 1 need not be first-order definable. This is due to a possible quantifier shift $L\exists$, $\exists L$; which reminds one of LM, ML, to which it is similar in syntactic (though not in semantic) structure.

$L\exists xAx \rightarrow \exists xLAx$ becomes a typically universally valid principle in those semantics for modal predicate logic which take *individual concepts* (i.e., functions f with domain W assigning, to each $w \in W$, an object $f(w) \in D_w$) for their basic individuals. It then expresses a form of the Axiom of Choice.

Finally, it may be noted that translation results like 9.10 may be proven here as well. To formulate an example, define *U-formulas* as formulas constructed from atomic formulas using $\forall$ and L only.

12.8 *Theorem.* For each sentence $\varphi \rightarrow \psi$ of modal predicate logic such that

(i) φ is constructed from U-formulas using $\top$, $\bot$, $\wedge$, $\vee$, M and $\exists$
(ii) ψ is constructed from atomic formulas, $\top$, $\bot$, $\wedge$, $\vee$, $\forall$, $\exists$, L and M,

an L_0^+-equivalent may be constructed effectively from $\varphi \rightarrow \psi$.

CHAPTER XIII

PRESERVATION CLASSES OF MODAL FORMULAS

The definition of $\bar{M}_1^{def}$ (9.14) involved being preserved in passing from a general frame satisfying a certain condition C to the underlying ("full") frame. A similar definition was given for the canonical modal logics (6.11). Formally, this gives rise to the concept of a "preservation class" defined by some condition C.

13.1 *Definition.* $M(C)$ is the set of all modal formulas φ such that, for all general frames $\langle \mathsf{F}, W\rangle$ satisfying C, if φ holds in $\langle \mathsf{F}, W\rangle$, then φ holds in F as well.

Probably the earliest example of such a class to be studied in the modal literature is S. K. Thomason's class "E" (cf. [79]); also known to K. Fine (cf. [24]) as the "natural logics". In this case, the condition C can be formulated as follows:

$$\langle \mathsf{F}, W\rangle \models \forall x \forall y(\forall P(Py \to Px) \to x = y)$$

(i.e., "identity of indiscernibles") and

$$\langle \mathsf{F}, W\rangle \models \forall x \forall y(\forall P(Py \to \exists z(Rxz \wedge Pz)) \to Rxy).$$

(Note that these two second-order sentences are true in all (full) frames.) A more involved choice of C yields Sahlqvist's "simple logics" of [66]. Clearly, the truly simplest case arises when C is the "empty condition".

13.2 *Lemma.* $M(-)$ is the set of modal formulas which are equivalent to a closed formula.

Proof. Obviously, any closed formula is in $M(-)$; because these formulas are invariant for the transition from $\langle \mathsf{F}, W \rangle$ to F. (Recall that closed formulas correspond — via ST — to L_0-formulas.) Conversely, suppose that $\varphi \in M(-)$. This may be reformulated in a more convenient form by noting that the smallest general frame $\langle \mathsf{F}, W \rangle$ on any given frame F has the set W consisting of exactly the sets $X \subseteq W$ of the form $X = \{w \in W \mid \mathsf{F} \models \psi[w]\}$; where ψ is a closed modal formula. For, thus, $\varphi \in M(-)$ if and only if $CS(\varphi) \models \varphi$; where $CS(\varphi)$ is the set of closed substitution instances of φ.
(I.e., for $\varphi = \varphi(p_1, ..., p_n)$,

$$CS(\varphi) = \{[\psi_1/p_1, ..., \psi_n/p_n]\varphi \mid \psi_1, ..., \psi_n \text{ are closed formulas}\}.)$$

(Cf. the proof of theorem 9.15.) Now, by compactness, $CS(\varphi) \models \varphi$ iff φ is implied by some finite conjunction of its closed substitution instances. It follows that φ is equivalent to this conjunction. QED.

One could start from the empty condition, building up hierarchies of classes $M(C)$ for ever stronger conditions C. E.g., the class E would be $M(C_1)$ in the hierarchy generated by the conditions

C_n: $\forall x \forall y (\forall P(Py \rightarrow \exists z(R^i xz \wedge Pz)) \rightarrow R^i xy)$ is true in $\langle \mathsf{F}, W \rangle$ for $0 \leqslant i \leqslant n$.

E was shown to be recursively enumerable and contained in $\bar{M}1$ by A. H. Lachlan in [47]. His first result may be generalized as follows.

13.3 *Lemma.* If C has the form "$\langle \mathsf{F}, W \rangle \models \alpha$" for some universal second-order sentence α, then $M(C)$ is recursively enumerable.

Proof. Lachlan's idea works in this more general case as well. It consists in describing the condition C in some suitable (many-sorted) first-order language: "If $\langle \mathsf{F}, W_1 \rangle$ is a general frame [three closure conditions, which may be spelt out] satisfying α as well as φ [this may be spelt out too], then any general frame $\langle \mathsf{F}, W_2 \rangle$ satisfies φ". QED.

If α holds on all frames, then $M(C) \subseteq \bar{M}1$. Lachlan proved this for the special case of E by applying the *interpolation theorem* to the above description in order to obtain an L_0-equivalent for φ. Here, a more general method is preferred, which is to be found, for the special case of (again) E, in R. I. Goldblatt [30].

13.4 *Theorem.* Let Γ be a set of sentences from monadic second-order logic which are true in all frames. For the condition $C =$ "$\langle \mathbf{F}, \mathbb{W} \rangle \models \Gamma$", $M(C) \subset \bar{M}1$.

Proof. In view of theorem 8.6, it suffices to show that any formula $\varphi \in M(C)$ is preserved under ultraproducts. Thus, consider a set $\{\mathbf{F}_i \mid i \in I\}$ of frames $\mathbf{F}_i$ such that $\mathbf{F}_i \models \varphi$, and an ultrafilter U on I. It is to be proven that $\mathbf{F} = \Pi_U \mathbf{F}_i \models \varphi$. To see this, consider the ultraproduct $\langle \mathbf{F}, \mathbb{W} \rangle$ of these same frames *considered as general frames* $\langle \mathbf{F}_i, P(W_i) \rangle$. (Cf. the discussion following theorem 4.12.) It follows from that theorem that $\langle \mathbf{F}, \mathbb{W} \rangle \models \varphi$; φ being true in all factors $\mathbf{F}_i$. Moreover, Γ, consisting of sentences true in all frames (and hence also true in all general frames $\langle \mathbf{F}_i, P(W_i) \rangle$) will be true in $\langle \mathbf{F}, \mathbb{W} \rangle$. But, then, since $\varphi \in M(C)$, $\mathbf{F} \models \varphi$. QED.

Theorem 13.4 implies that all classes $M(C_n)$, for C_n as defined above, are contained in $\bar{M}1$. No converse holds, however.

13.5 *Lemma.* $Lp \to MMp \in M1 - \bigcup_{n \in IN} M(C_n)$.

Proof. Theorem 9.8 yields the L_0-equivalent $\exists y(Rxy \wedge \exists z(Ryz \wedge Rxz))$ for $Lp \to MMp$. But, now consider the following general frame $\langle \mathbf{F}, \mathbb{W} \rangle$. $W = IN$, $R = \{\langle m, n \rangle \mid n = m + k$ for some odd number $k\}$, $\mathbb{W} = \{X \subseteq IN \mid X$ is finite or cofinite$\}$.

(i) $Lp \to MMp$ does not hold in $\mathbf{F}$, but
(ii) $Lp \to MMp$ does hold in $\langle \mathbf{F}, \mathbb{W} \rangle$, and
(iii) $\langle \mathbf{F}, \mathbb{W} \rangle$ satisfies all conditions C_n.

(i) is immediate. (ii) follows from the fact that, if $\langle \mathbf{F}, V \rangle \models Lp[n]$, then $V(p)$ is infinite and hence cofinite (being in $\mathbb{W}$). Therefore, for some m, $\{k \mid k \geqslant m\} \subseteq V(p)$, and, clearly, $\langle \mathbf{F}, V \rangle \models MMp[n]$. Finally, (iii) is proven by some honest calculation. QED.

More importantly, theorem 13.4 provides a new characterization of $\bar{M}1$. Let $UV2$ be the set of sentences from monadic second-order logic based upon L_0 (i.e., L_2) which are true in all frames.

13.6 *Theorem.* $\bar{M}1 = M(C)$ for the condition $C =$ "$\langle \mathsf{F}, W\rangle \models UV2$".

Proof. From 13.4, it follows that $M(C) \subseteq \bar{M}1$. For the converse direction, suppose that $\varphi \in \bar{M}1$. In other words, $\bar{E}(\varphi, \alpha)$ for some L_0-sentence α; or, in yet other words, $\varphi \leftrightarrow \alpha \in UV2$. Now, if $\langle \mathsf{F}, W\rangle \models \varphi$ and $\langle \mathsf{F}, W\rangle \models UV2$ then $\langle \mathsf{F}, W\rangle \models \varphi \leftrightarrow \alpha$ and, therefore, $\langle \mathsf{F}, W\rangle \models \alpha$. Since α is an L_0-sentence, it follows that $\mathsf{F} \models \alpha$ and, again by $\varphi \leftrightarrow \alpha$'s membership of $UV2$, $\mathsf{F} \models \varphi$. QED.

What has been established is that the first-order definable modal formulas are precisely the formulas which are invariant for the transition from a general frame "as much like a frame as possible" to the underlying frame. (Note that the proof of 13.6 uses only $UV2$-principles of the form $\forall x \forall P_1 ... \forall P_n ST(\varphi) \leftrightarrow \alpha$; which are equivalent to a conjunction of a $\Sigma^1_1(R)$-sentence and a $\Pi^1_1(R)$-sentence. Cf. chapter 17.)

Another interesting preservation class arises from the choice

$C =$ "W contains all one-element sets in W".

Note that this condition C implies all C_n's as defined previously. Moreover, $M(C) \subseteq \bar{M}_1^{def}$. (For, all singletons are among the sets which are L_0-definable in F using parameters; cf. definition 9.14.) On the other hand, $M(C) \neq \bar{M}_1^{def}$; as is shown by the formula $Lp \rightarrow MMp$ of 13.5. (To see this, just note that the set W defined there contains all singletons.) This class $M(C)$ is interesting for another reason too. The modal formula

$LM'\top \rightarrow L(L(Lp \rightarrow p) \rightarrow p)$

belongs to $M(C)$ (and hence to $\bar{M}_1^{def}$), but it is not *complete* in the sense of definition 6.4; as was proved in theorem 6.6. Thus, an obvious connection with Completeness Theory fails. As Sahlqvist has shown in [66], all modal formulas described in theorem 9.10 are complete. Since all these formulas are typically inside $\bar{M}_1^{def}$, it seems a reasonable conjecture that all formulas in $\bar{M}_1^{def}$ are complete. Unfortunately, the above formula constitutes a counter-example.

A preservation class which does consist entirely of complete formulas is CAN, defined as $M($"$\langle \mathsf{F}, W\rangle$ is descriptive"$)$. (Cf. definition 6.11.) But, again, we are disappointed: $CAN \nsubseteq \bar{M}1$; as was noted in chapter 6. Nevertheless, CAN is a very important class, which will be studied further in chapter 16.

To conclude, preservation classes give us interesting subsets of both $\bar{M}1$ and C (in the sense of definition 6.4). Even $\bar{M}1$ itself is characterizable as a preservation class (13.6). For C, no similar result is known, however. An even more interesting set (from the point of view of Correspondence Theory) is $\bar{M}1 \cap C$. Because of chapter 6 (8), $\bar{M}1 \cap C = \bar{M}1 \cap CAN$; whence it may be written as an intersection of two preservation classes. Therefore, it is a preservation class itself, viz. M("$\langle \mathsf{F}, W \rangle \models UV2$ or $\langle \mathsf{F}, W \rangle$ is descriptive"). Can a more natural and elegant condition be formulated which generates $\bar{M}1 \cap C$ as a preservation class?

PART III

Modal Definability

CHAPTER XIV

MODALLY DEFINABLE ELEMENTARY CLASSES OF FRAMES

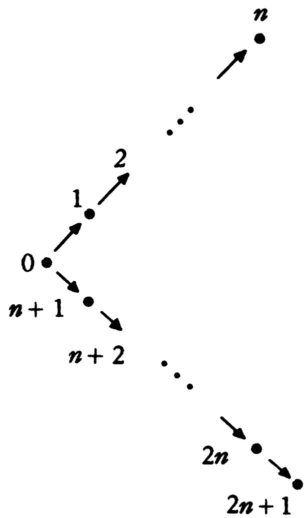
In chapter 3, the relations E of *local equivalence* and $\bar{E}$ of *global equivalence* between modal formulas and L_0-formulas were introduced (definition 3.11). This definition can be extended to *sets* Σ of modal formulas and sets Δ of L_0-formulas. There are also "mixed" cases, yielding four combinations in all: (i)φ, α, (ii)φ, Δ, (iii)Σ, α and (iv)Σ, Δ. It was shown in chapter 3 that, if a modal formula is equivalent (in either sense) to a set Δ of L_0-formulas, then it is equivalent to some finite subset of Δ; i.e., to a single L_0-formula (viz. the conjunction of that finite set). Thus, in a sense, case (ii) reduces to case (i). On the other hand, case (iv) is irreducible in that Σ and Δ exist which are equivalent, but no single modal formula or L_0-sentence defines $FR(\Sigma) = FR(\Delta)$. A global example of this is provided by (cf.[6]):

$$\Sigma = \{M^iL\bot \rightarrow L^iL\bot \mid i \geqslant 1\},$$
$$\Delta = \{\forall x(\exists y(R^ixy \neg\exists zRyz) \rightarrow \forall y(R^ixy \rightarrow \neg \exists zRyz)) \mid i \geqslant 1\}.$$

It is obvious, using lemma 9.6, that $\bar{E}(\Sigma, \Delta)$. Moreover, no L_0-sentence α defines $FR(\Delta)$. For, in that case, $\Delta \models \alpha$, whence — by compactness — $\Delta_0 \models \alpha$ for some finite $\Delta_0 \subseteq \Delta$. Since, on the other hand, $\alpha \models \delta$ for each $\delta \in \Delta$, it would follow that $\Delta_0 \models \delta$ for each $\delta \in \Delta$. This, however, can be refuted by considering the frames $\mathbf{F}_n = \langle W_n, R_n\rangle$, where, for $n \geqslant 1$,

$$W_n = \{0, 1, ..., n, n + 1, ..., 2n + 1\}, \quad \text{and}$$

$$R_n = \{\langle i, i+1\rangle \mid 0 \leqslant i \leqslant n-1\} \cup \{\langle 0, n+1\rangle\} \cup \{\langle i, i+1\rangle \mid n+1 \leqslant i \leqslant 2n\}.$$



It is easy to check that

$$\mathbf{F}_n \models \forall x(\exists y(R^i xy \wedge \neg\exists z Ryz) \rightarrow \forall y(R^i xy \rightarrow \neg\exists z Ryz))$$

for each i $(1 \leqslant i < n)$. But

$$\mathbf{F}_n \not\models \forall x(\exists y(R^n xy \wedge \neg\exists z Ryz) \rightarrow \forall y(R^n xy \rightarrow \neg\exists z Ryz)):$$

look at 0. Clearly, then, no finite subset of Δ can imply all of Δ. Finally, since $FR(\Delta)$ is not definable by means of a single L_0-sentence, it is not definable by means of a single modal formula (instead of the set Σ); because, if it were, then $FR(\Delta)$ would be L_0-elementary after all.

The only remaining case is Σ, α. Unfortunately, it has not been possible either to reduce this case to case (i) or to show that it is irreducible. Thus, a general question remains:

"Are there sets of modal formulas defining an L_0-elementary class of frames which is not definable by means of a single modal formula?"

For a special case, (iii) reduces to (i): viz. that of an L_0-sentence equivalent to a set Σ which is ***natural*** in the sense of K.Fine [24](cf. chapter 13); i.e., which is preserved in passing from a general frame in which

$$\forall x \forall y(\forall P(Py \rightarrow Px) \rightarrow x = y) \quad \text{and}$$
$$\forall x \forall y(\forall P(Py \rightarrow \exists z(Rxz \wedge Pz)) \rightarrow Rxy)$$

hold, to the underlying full frame.

14.1 *Lemma.* If $\bar{E}(\Sigma, \alpha)$ for a natural set Σ and an L_0-sentence α, then $\bar{E}(\varphi, \alpha)$ for some modal formula φ.

Proof. This lemma may be established using the algebraic method of chapter 4. One considers the set Σ^* of polynomial identities corresponding to formulas in Σ and shows that the class $\{\mathbf{A} \mid \mathbf{A} \models \Sigma^*\}$ of modal algebras in which these hold is definable by means of a *finite* number of such identities. The conjunction of the corresponding modal formulas will then define $FR(\Sigma)$.

To prove that $\{\mathbf{A} \mid \mathbf{A} \models \Sigma^*\}$ is definable using one single polynomial identity, it suffices to show that the complementary class $\{\mathbf{A} \mid \mathbf{A} \not\models \Sigma^*\}$ is closed under ultraproducts. For, then, by Keisler's characterization of elementary classes, $\{\mathbf{A} \mid \mathbf{A} \models \Sigma^*\}$ will be definable by means of a single sentence α (itself not necessarily a polynomial identity). By compactness, α, being a logical consequence of Σ^*, follows from some finite set $\Sigma_0 \subseteq \Sigma^*$. Conversely, α implies each identity in Σ^*. In other words, Σ_0 defines $\{\mathbf{A} \mid \mathbf{A} \models \Sigma^*\}$

Let, then, $\mathbf{A}_i \not\models \Sigma^*$ for each $i \in I$, and let U be any ultrafilter on I. It is to be shown that $\Pi_U \mathbf{A}_i \not\models \Sigma^*$. Recall the *Stone Representation* of chapter 4: each $\mathbf{A}_i$ is isomorphic to $SR(\mathbf{A}_i)^+$; where $SR(\mathbf{A}_i)$ is a general frame $\langle \mathbf{F}_i, W_i \rangle$. (Cf. definition 4.5.) Moreover, $SR(\mathbf{A}_i)$ is *descriptive*; which means that — amongst others — the two above-mentioned second-order sentences hold in it. Note that, furthermore, $\langle \mathbf{F}_i, W_i \rangle \not\models \Sigma$, whence $\mathbf{F}_i \not\models \Sigma$ and, therefore, $\mathbf{F}_i \models \neg\alpha$.

Now, consider the ultraproduct $\langle \mathbf{F}, W \rangle = \Pi_U\{\langle \mathbf{F}_i, W_i \rangle \mid i \in I\}$ (cf. definition 4.11). By the theorem of Łoś, both $\neg\alpha$ and the above-mentioned second-order sentences hold in $\langle \mathbf{F}, W \rangle$. But, then, $\langle \mathbf{F}, W \rangle \not\models \Sigma$; for, otherwise, $\mathbf{F} \models \Sigma$ (Σ being natural) and $\mathbf{F} \models \alpha$: a contradiction.

Finally, to see that $\Pi_U \mathbf{A}_i \not\models \Sigma^*$, it suffices to note that $\Pi_U \mathbf{A}_i = \Pi_U \langle \mathbf{F}_i, W_i \rangle^+ \cong (\Pi_U \langle \mathbf{F}_i, W_i \rangle)^+$ (by lemma 4.13). QED.

Unfortunately, lemma 14.1 does not solve our question in its entirety. If $\bar{E}(\Sigma, \alpha)$, then Σ need not be natural. Indeed, $FR(\alpha)$ need not be definable by any natural Σ. E.g., recall the fact (cf. chapter 7) that $\bar{E}(\varphi, \alpha)$ for

$$\varphi = (Lp \to LLp) \wedge (LMp \to MLp)$$
$$\alpha = \forall x \forall y (Rxy \to \forall z (Ryz \to Rxz)) \wedge \forall x \exists y (Rxy \wedge \forall z (Ryz \to z = y)).$$

Suppose that also $\bar{E}(\Sigma, \alpha)$ for some natural Σ. Now, φ is *complete* in the sense of definition 9.4 (cf. Segerberg [68]). Therefore, since $\varphi \models_f \sigma$ for each $\sigma \in \Sigma$, it holds that $\varphi \models_{gf} \sigma$ for each $\sigma \in \Sigma$. So, φ must be natural. For, suppose that it holds on some general frame satisfying the relevant condition. Σ will hold on that general frame as well, and hence — being natural — it holds on the underlying frame. Clearly, then, α holds there and, therefore, φ holds there too. But, now there is a contradiction; for it follows from the example given in the proof of theorem 9.17 that φ is not natural.

The only thing which can be said for certain about Σ's satisfying $\bar{E}(\Sigma, \alpha)$ for some L_0-sentence α is that $FR(\Sigma)$ is also defined by the *canonical* set (cf. definitions 6.11, 6.8) $C(\Sigma)$. (That $C(\Sigma)$ is canonical follows from its being both first-order definable and complete; cf. corollary 16.7). The example of the above φ shows, then, that not all canonical sets Σ are natural (cf. Fine [24]).

In part II, the "modal side" of the equivalences E and $\bar{E}$ was studied, in the form of $M1$ and $\bar{M}1$. Now, the "predicate-logical side" will be the subject of investigation. Recall the relevant definitions (3.11):

$P1 = \{\alpha \mid \alpha$ is an L_0-formula with one free variable such that, for some modal formula φ, $E(\varphi, \alpha)\}$

$\bar{P}1 = \{\alpha \mid \alpha$ is an L_0-sentence such that $\bar{E}(\varphi, \alpha)$ holds for some modal formula $\varphi\}$.

The local notion is less interesting from the point of view of L_0; and $P1$ will, therefore, receive comparatively little attention. The main result is a semantic characterization of the global class $\bar{P}1$, due to R. I. Goldblatt and S. K. Thomason (cf. [31]); cf. theorem 14.7 below. We start with some simple observations about $P1$, however.

14.2 *Lemma.* If α and β are L_0-formulas with one and the same free variable (x), then

(i) if $\alpha \in P1$ and $\beta \in P1$, then $\alpha \wedge \beta \in P1$
(ii) if $\alpha \in P1$ and $\beta \in P1$, then $\alpha \vee \beta \in P1$
(iii) if $\alpha \in P1$, then $\forall x(Ryx \rightarrow \alpha) \in P1$, provided that y is distinct from x.

Proof. (i) follows from lemma 9.1; and so does (ii): if $E(\varphi, \alpha)$ and (ψ, β), then change the proposition letters in φ to new ones not occurring in ψ. This yields φ', and $E(\varphi' \vee \psi, \alpha \vee \beta)$. Finally, if $E(\varphi, \alpha)$, then $E(L\varphi, \forall x(Ryx \rightarrow \alpha))$ (cf. corollary 9.2). QED.

14.3 *Lemma.* *P1* is not closed under negation; *P1* is not closed under restricted existential quantification.

Proof. $Rxx \in P1$, because $E(\mathrm{L}p \rightarrow p, Rxx)$. On the other hand, $\neg Rxx \notin P1$. This follows from the example of $\langle IN, <\rangle$ with the p-morphism f onto $\langle\{0\}, \{\langle 0, 0\rangle\}\rangle$ defined by $f(n) = 0$ for all $n \in IN$. For, $\langle IN, <\rangle \models \neg Rxx[0]$; but $\langle\{0\}, \{\langle 0, 0\rangle\} \not\models \neg Rxx[f(0)]$ (and apply corollary 2.18).

The proof (following corollary 2.31) that $\forall x \exists y(Rxy \wedge Ryy)$ is not in $\bar{P}1$ may be modified in an obvious way to show that $\exists y(Rxy \wedge Ryy)$ is not in *P1*. This proves the second assertion. QED.

Using the proofs of theorems 9.8 and 9.10, an important subclass of *P1* may be described constructively. Here only a very partial result will be proven. First, an auxiliary notion is needed.

14.4 *Definition.* A $\forall$-*formula* is an L_0-formula with one free variable, which is of the form $U\alpha$, where U is a (possibly empty) sequence of restricted universal quantifiers and α is an L_0-formula in which only atomic formulas and the Boolean operators $\wedge$ and $\vee$ occur.

Many relational conditions occurring in the literature are expressible using $\forall$-formulas; e.g., reflexivity (Rxx), symmetry ($\forall y(Rxy \rightarrow Ryx)$) and transitivity ($\forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))$). But also the oft-mentioned property of having no more than a given number n of R-incomparable R-successors is so definable:

$$\forall y_1(Rxy_1 \rightarrow \ldots \rightarrow \forall y_n(Rxy_n \rightarrow \forall y_{n+1}(Rxy_{n+1} \rightarrow \sum_{1 \leqslant i \neq j \leqslant n+1} y_i = y_j \vee Ry_iy_j)))\ldots)).$$

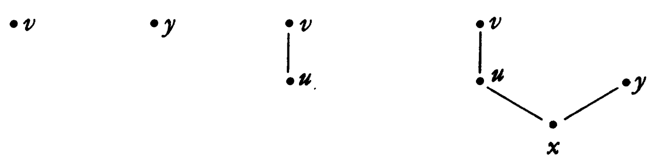
14.5 *Lemma.* Each $\forall$-formula has a modal equivalent which is constructively obtainable from it.

Proof. Consider any $\forall$-formula $U\alpha$. Using the propositional distributive laws, write α as a conjuction $\prod_{i=1}^{n} \alpha_i$ of disjunctions α_i of atomic formulas.

Since $U\alpha$ is equivalent to $\prod_{i=1}^{n} U\alpha_i$, it suffices to consider the conjuncts $U\alpha_i$, by lemma 14.2. Now, rewrite $U\alpha_i$ to a formula β_i of the form "$\neg$ -sequence of restricted existential quantifiers - conjunction of negated atomic formulas". Remove repetitions from this conjunction, and also drop one of each pair of the form $\neg\, x = y, \neg\, y = x$ occurring in it. Finally, provide each quantifier with a unique bound variable.

A tree $T(y)$ is constructed inductively for each bound variable y occurring in β_i. If no restricted existential quantifiers of the form $\exists z(Ryz \wedge$ occur in β_i, then $T(y)$ consists of a single node y. Otherwise, $T(y)$ is constructed from the trees $T(z_1), \ldots, T(z_m)$, where $z_1, \ldots, z_m$ are the variables for which $\exists z(Ryz \wedge$ occurs in β_i, by joining their top nodes to a new top node y. An example is provided by $\forall y(Rxy \rightarrow \forall u(Rxu \rightarrow \forall v(Ruv \rightarrow Ryv)))$. Rewriting this as $\neg\exists y(Rxy \wedge \exists u(Rxu \wedge \exists v(Ruv \wedge \neg Ryv)))$ yields the trees

$T(v)$: $\qquad$ $T(y)$: $\qquad$ $T(u)$: $\qquad$ and $T(x)$:



For each node y in the tree $T(x)$ (where x is the one free variable in $U\alpha_i$), a modal formula $\mathbf{mod}(y)$ is defined inductively as the conjuction of the formulas

(i)	$M\,\mathbf{mod}(z)$	for each immediate descendant z of y,
(ii)	Lp_{yz}	for each formula $\neg Ryz$ occurring in the propositional matrix of β,
(iii)	$\neg p_{zy}$	for each formula $\neg Rzy$ occurring in the propositional matrix of β_i,
(iv)	q_{yz}	for each formula $\neg y = z$ occurring in the propositional matrix of β_i,
(v)	$\neg q_{zy}$	for each formula $\neg z = y$ occurring in the propositional matrix of β_i,

or

(vi)	$\top$	if the above conjunction is empty.

(The subscripts of the proposition letters p and q indicate that distinct proposition letters are to be chosen for these respective roles.) E.g., in the above example, $\mathrm{mod}(y) = Lp_{yv}$, $\mathrm{mod}(v) = \neg p_{yv}$, $\mathrm{mod}(u) = M\neg p_{yv}$, $\mathrm{mod}(x) = MM\neg p_{yv} \wedge MLp_{yv}$.

The modal formula equivalent to $U\alpha_i$ is $\neg\mathrm{mod}(x)$. This is easily shown by noting that, for all frames $\mathbf{F} = \langle W, R\rangle$, and all $w \in W$, $\mathbf{F} \not\models U\alpha_i[w]$ iff, for some valuation V on $\mathbf{F}$, $\langle \mathbf{F}, V\rangle \models \mathrm{mod}(x)[w]$. Thus, e.g., the above L_0-formula is equivalent to $\neg(MM\neg p_{yv} \wedge MLp_{yv})$, i.e., to $MLp_{yv} \rightarrow LLp_{yv}$. QED.

Next, the class $\bar{P}1$ will be studied, which lends itself more elegantly to model-theoretic investigations. Note that $P1 \subseteq \bar{P}1$. (There seems to be no very natural formulation for a converse result to the effect that "$\bar{P}1 \subseteq P1$".) Like $P1$, $\bar{P}1$ is closed under conjunction and universal quantification. It is not closed under disjunction, hovewer. Consider, e.g., $\forall xRxx$ and $\forall x\forall y(Rxy \rightarrow Ryx)$; which are both in $\bar{P}1$. Their disjunction $\forall xRxx \vee \forall x\forall y(Rxy \rightarrow Ryx)$ is not in $\bar{P}1$; not being preserved under disjoint unions. (This sentence holds, e.g., in both $\langle\{0\}, \emptyset\rangle$ and $\langle\{0, 1\}, \{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 1, 1\rangle\}\rangle$, but not in their disjoint union.)

To state the semantic characterization of $\bar{P}1$, the following concept is useful.

14.6 *Definition.* A class $\mathbf{K}$ of frames is *modally definable* if some set Σ of modal formulas exists such that $\mathbf{K} = FR(\Sigma)$.

14.7 *Theorem* (R. I. Goldblatt & S. K. Thomason [31]; cf. 2.28). If a class $\mathbf{K}$ of frames is closed under (L_0-)elementary equivalence, then $\mathbf{K}$ is modally definable if and only if $\mathbf{K}$ is closed under generated subframes. disjoint unions and p-morphic images, while its complement ${}^c\mathbf{K}$ is closed under ultrafilter extensions.

Proof. If $\mathbf{K}$ is modally definable, then $\mathbf{K}$ and ${}^c\mathbf{K}$ satisfy the mentioned closure properties; by 2.12, 2.15, 2.18 and 2.26, respectively. If $\mathbf{K}$ and ${}^c\mathbf{K}$ satisfy these closure conditions (which imply closure of $\mathbf{K}$ under ultrafilter extensions as well; by theorem 8.9), then the conditions of theorem 16.5 below are satisfied, and $\mathbf{K}$ turns out to be definable by means of a set of modal formulas which is even canonical. QED.

For the class $\bar{P}1'$ of L_0-sentences defined by a set of modal formulas, theorem 14.7 implies

14.8 *Corollary.* For any L_0-sentence α, $\alpha \in \bar{P}1'$ if and only if α is preserved under generated subframes, disjoint unions and p-morphic images, while $\neg\alpha$ is preserved under ultrafilter extensions.

For $\bar{P}1$ "sec", a semantic characterization may be found in Goldblatt [30] (theorem 20.10); but the additional notion involved ("completed ultra-product") is not as (elegant and) natural as the ones appearing above. Therefore, it has been omitted here.

Corollary 14.8 yields a model-theoretic proof for the global version of lemma 14.5. Let α be any L_0-sentence of the form $\forall x\beta$; where β is an $\forall$-formula in the sense of definition 14.4. It will be shown in chapter 15 that such formulas α are preserved under generated subframes, disjoint unions and p-morphic images. Moreover, $\neg\alpha$, being equivalent to an existential L_0-sentence, is preserved under ultrafilter extensions. (Use the fact that $ue(\mathbf{F})$ is (isomorphic to) an extension of $\mathbf{F}$ for each frame $\mathbf{F}$; cf. lemma 2.27.) This proof is highly non-constructive, of course.

CHAPTER XV

PRESERVATION RESULTS FOR FIRST-ORDER FORMULAS

In this chapter, some preservation results will be proven for L_0-formulas with respect to the main semantic operations on frames introduced up to now. This yields syntactic information about $P1$ and $\bar{P}1$, as may be seen from corollaries 15.5 and 15.16 below. In fact, one could expect a complete characterization. No such result has been found, however. *Necessary* syntactic conditions for membership of $P1$ or $\bar{P}1$ are available, necessary *and sufficient* conditions are not. Indeed, the question if $P1$ and $\bar{P}1$ are at least recursively enumerable is still open. It has been conjectured by M. H. Löb that $M1(\bar{M}1)$ and $P1(\bar{P}1)$ are recursive in each other. This would transfer the doubts about the arithmetical definability of $\bar{M}1$ (expressed in chapter 7) to $\bar{P}1$. On the other hand, there is no suggestive result to be expected from the more general (and tractable) case of universal second-order sentences (cf. theorem 17.10). Indeed, there is no non-trivial generalization of $\bar{P}1$ to this field.

The proofs of the following results will mainly consist in *elementary chain* constructions like in the proof of theorem 3.9. *Saturated structures* could be used as well (cf. chapter 5 of Chang & Keisler [17]), but no gain in insight would compensate for the additional technicalities involved in that method.

The first notions and results are concerned with $P1$.

15.1 *Definition* (cf. Lemma 11.4). An L_0-formula $\alpha = \alpha(x_1,..., x_n)$ is *invariant for generated subframes* if, for all frames $\mathsf{F}_1(= \langle W_1,R_1\rangle)$ and F_2 such that $\mathsf{F}_1 \underset{\rightarrow}{\subseteq} \mathsf{F}_2$ and for all $w_1,..., w_n \in W_1$,

$$\mathsf{F}_1 \models \alpha[w_1, ..., w_n] \text{ iff } \mathsf{F}_2 \models \alpha[w_1, ..., w_n].$$

15.2 *Definition.* An L_0-formula $\alpha = \alpha(x_1,..., x_n)$ is ***preserved under p-morphisms*** if, for all frames $\mathsf{F}_1(=\langle W_1,R_1\rangle)$ and F_2, all p-morphisms f from F_1 onto F_2, and all $w_1, ..., w_n \in W_1$,

$$\mathsf{F}_1 \models \alpha[w_1, ..., w_n] \text{ only if } \mathsf{F}_2 \models \alpha[f(w_1), ..., f(w_n)].$$

15.3 *Definition.* Let L be a first-order language containing the fixed binary predicate constant R. The ***restricted positive formulas*** of L are the L-formulas constructed from $\perp$ and atomic formulas (with or without individual constants) using $\wedge$, $\vee$, restricted universal quantifiers $\forall y(Rty\rightarrow$ and restricted existential quantifiers $\exists y(Rty\wedge$ (where t is an individual constant or variable distinct from y).

15.4 *Theorem.* Any L_0-formula is invariant for generated subframes and preserved under p-morphisms if and only if it is logically equivalent to a restricted positive L_0-formula with the same free variables.

Proof. The direction from left to right follows from the next two statements, which are easily proven using induction on the complexity of formulas. Any restricted (positive) formula of L_0 is invariant for generated subframes. Any restricted positive formula of L_0 is preserved under p-morphisms.

Now, let the L_0-formula $\alpha = \alpha(x_1,..., x_n)$ be invariant for generated subframes and preserved under p-morphisms. Define $RP(\alpha)$ as

$$\{\beta|\beta = \beta(x_1,..., x_n) \text{ is a restricted positive } L_0\text{-formula such that } \alpha \models \beta\}.$$

It will be shown that $RP(\alpha) \models \alpha$. After that, the compactness theorem yields finitely many $\beta_1, ..., \beta_m \in RP(\alpha)$ implying α; whence α is logically equivalent to $\beta_1 \wedge ... \wedge \beta_m$.

Suppose, then, that $\mathsf{F}_1^! \models RP(\alpha)[w_1, ..., w_n]$. New individual constants $\mathbf{w}_1, ..., \mathbf{w}_n$ are added to L_0 to obtain a first-order language L_1; and $\mathsf{F}_1^!$ is expanded to an L_1-structure F_1 by interpreting each $\mathbf{w}_i$ as w_i. (**N.B.**: In the remainder of this chapter, "L_1" does not denote the first-order language of part I, but some purely ad-hoc first-order language needed in the relevant proofs.)

Each finite subset of $\Sigma = \{[\mathbf{w}_1/x_1, ..., \mathbf{w}_n/x_n]\alpha\} \cup \{\neg\beta \mid \beta \text{ is a restricted positive sentence of } L_1 \text{ such that } \mathsf{F}_1 \models \neg\beta\}$ has a model. Otherwise, it would follow that $[\mathbf{w}_1/x_1, ..., \mathbf{w}_n/x_n]\alpha \models \neg(\neg\beta_1 \wedge ... \wedge \neg\beta_k)$ for some

$\neg\beta_1, \ldots, \neg\beta_k$ as described; whence $[\mathbf{w}_1/x_1, \ldots, \mathbf{w}_n/x_n]\alpha \models \beta_1 \vee \ldots \vee \beta_k$. But, $\beta_1 \vee \ldots \vee \beta_k$ is a restricted positive L_1-sentence. Moreover, $\alpha \models [x_1/\mathbf{w}_1, \ldots, x_n/\mathbf{w}_n](\beta_1 \vee \ldots \vee \beta_k)$ — provided that bound variables x_i are changed in $\beta_1 \vee \ldots \vee \beta_k$ whenever necessary. (Such small details concerning *substitutability* of variables will be neglected henceforth.) In other words,

$$\beta = [x_1/\mathbf{w}_1, \ldots, x_n/\mathbf{w}_n](\beta_1 \vee \ldots \vee \beta_k) \in RP(\alpha);$$

whence $\mathbf{F}_1^1 \models \beta[w_1, \ldots, w_n]$ and $\mathbf{F}_1 \models \beta_1 \vee \ldots \vee \beta_k$, contradicting the fact that $\mathbf{F}_1 \models \neg\beta_1, \ldots, \mathbf{F}_1 \models \neg\beta_k$. It follows, by the compactness theorem, that Σ has a model, say $\mathbf{G}_1$. (From now on, capital letters "**F**" and "**G**", possibly with subscripts and/or superscripts, will be used for frames; possibly with distinguished elements.) This yields the following situation:

$$\begin{tikzcd}
\textit{frames:} & \mathbf{F}_1^1 \,, & \mathbf{F}_1 \\
& & \mathbf{G}_1, \arrow[u] \\
\textit{languages:} & L_0 \,, & L_1
\end{tikzcd}$$

where (i) $\mathbf{G}_1 \models [\mathbf{w}_1/x_1, \ldots, \mathbf{w}_n/x_n]\alpha$
and (ii) $\mathbf{G}_1 - RP(L_1) - \mathbf{F}_1$; which notation abbreviates: "all restricted positive sentences of L_1 which are true in $\mathbf{G}_1$ are also true in $\mathbf{F}_1$".

Starting from this situation, elementary chains $\mathbf{F}_1, \mathbf{F}_2, \ldots$ and $\mathbf{G}_1, \mathbf{G}_2, \ldots$ will be constructed, together with ever increasing first-order languages L_1, L_2, ... The general method is as follows.

Let a language L_n be given, and L_n-structures $\mathbf{F}_n$ and $\mathbf{G}_n$ such that $\mathbf{G}_n - RP(L_n) - \mathbf{F}_n$. For each individual constant c in L_n and each w in the domain of $\mathbf{G}_n$ such that $\mathbf{G}_n \models Rcx[w]$, add a *new* individual constant $\mathbf{w}$ to L_n to obtain a language L_n^1. Expand $\mathbf{G}_n$ to an L_n^1-structure $\mathbf{G}_n^1$ by interpreting each $\mathbf{w}$ as w.

Each finite subset of $\Delta = \{\beta \mid \beta$ is a restricted positive sentence of L_n^1 such that $\mathbf{G}_n^1 \models \beta\}$ has a model which is an expansion of $\mathbf{F}_n$. For, let $\beta_1, \ldots, \beta_k \in \Delta$, containing (in all) the constants $\mathbf{w}_1, \ldots, \mathbf{w}_s$ from $L_n^1 - L_n$. Constants $c_1, \ldots, c_s$ of L_n exist such that, for variables $x_1, \ldots, x_s$ not occurring in $\beta_1, \ldots, \beta_k$,

$$\mathbf{G}_n \models Rc_1x_1 \wedge \ldots \wedge Rc_sx_s \wedge [x_1/\mathbf{w}_1, \ldots, x_s/\mathbf{w}_s](\beta_1 \wedge \ldots \wedge \beta_k)[w_1, \ldots, w_s].$$

Therefore,

$$\mathbf{G}_n \models \exists x_1(Rc_1x_1 \wedge \ldots \wedge \exists x_s(Rc_sx_s \wedge [x_1/\mathbf{w}_1, \ldots, x_s/\mathbf{w}_s](\beta_1 \wedge \ldots \wedge \beta_k).$$

This formula is a restricted positive sentence of $L_n(!)$ and so it holds in $\mathbf{F}_n$; by virtue of $\mathbf{G}_n - RP(L_n) - \mathbf{F}_n$. From this the claim follows easily. A standard model-theoretic argument now establishes the existence of a model $\mathbf{F}^1_n$ for Δ such that

(i) $\mathbf{F}^1_n$ is an L^1_n-structure

(ii) $\mathbf{F}_n \prec_{L_n} \mathbf{F}^1_n$; i.e., $\mathbf{F}_n$ is an L_n-elementary substructure of $\mathbf{F}^1_n$,

and (iii) $\mathbf{G}^1_n - RP(L^1_n) - \mathbf{F}^1_n$.

Picture this as

$$\begin{tikzcd}
\textit{frames:} & \mathbf{F}_n & \prec & \mathbf{F}^1_n, \\
& \mathbf{G}_n, \arrow[u] & \mathbf{G}^1_n \arrow[ur] & \\
\textit{languages:} & L_n, & L^1_n, & L^1_n,
\end{tikzcd}$$

Next comes the dual construction step.

For each individual constant c of L^1_n and each w in the domain of $\mathbf{F}^1_n$ such that $\mathbf{F}^1_n \models Rcx[w]$, add a *new* individual constant k_{cw} to L^1_n to obtain L_{n+1}. Expand $\mathbf{F}^1_n$ to an L_{n+1}-structure $\mathbf{F}_{n+1}$ by interpreting each k_{cw} as w. (The use of k_{cw} instead of just $\mathbf{w}$, as above, will become clear presently.)

Each finite subset of $\Gamma = \{\neg\beta | \beta$ is a restricted positive sentence of L_{n+1} such that $\mathbf{F}_{n+1} \models \neg\beta\} \cup \{Rck_{cw} \mid k_{cw}$ is a constant of $L_{n+1} - L^1_n\}$ has a model which is an expansion of $\mathbf{G}^1_n$. To see this, let $\neg\beta_1, \ldots, \neg\beta_k \in \Gamma$ and consider any finite set $Rc_1k_{c_1w_1}, \ldots, Rc_sk_{c_sw_s}$. If $\neg\beta_1, \ldots, \neg\beta_k$ contain more constants k_{cw} from $L_{n+1} - L^1_n$ than $k_{c_1w_1}, \ldots, k_{c_sw_s}$, then the relevant Rck_{cw}'s may be added. Thus, it may be supposed that all constants from $L_{n+1} - L_n$ occurring in $\neg\beta_1, \ldots, \neg\beta_k$ are among $k_{c_1w_1}, \ldots, k_{c_sw_s}$. Now, suppose that $\{\neg\beta_1, \ldots, \neg\beta_k, Rc_1k_{c_1w_1}, \ldots, Rc_sk_{c_sw_s}\}$ is *not* satisfiable in any expansion of $\mathbf{G}^1_n$. It follows, for any sequence of variables $x_1, \ldots, x_s$ not occurring in $\neg\beta_1, \ldots, \neg\beta_k$, that

$$\mathbf{G}^1_n \models \forall x_1(Rc_1x_1 \rightarrow \ldots \rightarrow \forall x_s(Rc_sx_s \rightarrow [x_1/k_{c_1w_1}, \ldots, x_s/k_{c_sw_s}](\beta_1 \vee \ldots \vee \beta_k))\ldots).$$

Moreover, this restricted positive $L_n^1(!)$-sentence holds in $\mathbf{F}_n^1$; because $\mathbf{G}_n^1 - RP(L_n^1) - \mathbf{F}_n^1$. But this contradicts the fact that $\mathbf{F}_n^1 \models Rc_1x_1 \wedge \ldots \wedge Rc_sx_s \wedge [x_1/k_{c_1w_1}, .., x_s/k_{c_sw_s}](\neg\beta_1 \wedge \ldots \wedge \neg\beta_n)[w_1, \ldots, w_s]$. (This follows from the fact that each sentence $Rc_ik_{c_iw_i}$ holds in $\mathbf{F}_{n+1}$; i.e., $\mathbf{F}_n^1 \models Rc_ix_i\,[w_i]$, and each sentence $\neg\beta$ holds in $\mathbf{F}_{n+1}$ by assumption.)

Two remarks should be made at this point. First, the constants k_{cw} are used, instead of just $\mathbf{w}$, to avoid the following situation. For $\{\neg\beta(\mathbf{w}), Rc_1\mathbf{w}, Rc_2\mathbf{w}\}$, the above argument would yield the formula $\forall y(Rc_1y \rightarrow (Rc_2y \rightarrow \beta(y)))$, which is not restricted positive. Put differently, one cannot talk about "common R-successors": $\{\neg\beta(k_{c_1w}), \neg\beta(k_{c_2w}), Rc_1k_{c_1w}, Rc_2k_{c_2w}\}$ is the best approximation. (Cf. the remark following the proof of theorem 3.9.)' Secondly, if, in the above argument, the limiting case is considered without any sentences $\neg\beta$, then it becomes necessary to use $\perp$. To see that $\{Rc_1k_{c_1w}, Rc_2k_{c_2w}\}$ is satisfiable in an expansion of $\mathbf{G}_n^1$, one supposes that $\mathbf{G}_n^1 \models \forall x_1(Rc_1x_1 \rightarrow \ldots \rightarrow \forall x_s(Rc_sx_s \rightarrow \perp)\ldots)$ etc. Here is where the $\perp$ of definition 15.3 is needed essentially. In fact, all one needs is at least one restricted positive sentence β of L_{n+1} such that $\mathbf{F}_{n+1} \models \neg\beta$. $\perp$ is such a sentence, and in certain cases — e.g., when $\mathbf{F}_{n+1} = \langle\{0\}, \{\langle 0, 0\rangle\}\rangle$ — it may be the only one.

Again, a standard model-theoretic argument establishes the existence of a model $\mathbf{G}_{n+1}$ for Γ satisfying

(i) $\mathbf{G}_{n+1}$ is an L_{n+1}-structure
(ii) $\mathbf{G}_n^1 \prec_{L_n^1} \mathbf{G}_{n+1}$, and
(iii) $\mathbf{G}_{n+1} - RP(L_{n+1}) - \mathbf{F}_{n+1}$.

This may be pictured as follows:

$$\begin{tikzcd}[column sep=small]
\textit{frames} & \mathbf{F}_n & \prec & \mathbf{F}_n^1, & \mathbf{F}_{n+1} \\
& \mathbf{G}_n, \arrow[u] & \mathbf{G}_n^1 \arrow[ur] & \prec & \mathbf{G}_{n+1}, \\
\textit{languages:} & L_n, & L_n^1, & L_n^1, & L_{n+1},
\end{tikzcd}$$

From this elaborate exposition, il will be clear how the promised elementary chains $\mathbf{F}_1, \mathbf{F}_2, \ldots$ and $\mathbf{G}_1, \mathbf{G}_2, \ldots$ are constructed. Now, several applications of Tarski's fundamental theorem on elementary chains, in combination with the original assumptions on α, will yield the required conclusion that $\mathbf{F}_1^! \models \alpha[w_1, \ldots, w_n]$.

First, because $\mathbf{G}_1 \models [\mathrm{w}_1/x_1, \ldots, \mathrm{w}_n/x_n]\alpha$, this L_1-sentence holds in the limit $\mathbf{G}$ of the chain $\mathbf{G}_1, \mathbf{G}_2, \ldots$. Next, consider the smallest generated subframe $TC(\mathbf{G}, \mathrm{w}_1^{\mathbf{G}}, \ldots, \mathrm{w}_n^{\mathbf{G}})$ of $\mathbf{G}$ containing $\mathrm{w}_1^{\mathbf{G}}, \ldots, \mathrm{w}_n^{\mathbf{G}}$. By the above construction, this subframe is exactly the substructure of $\mathbf{G}$ whose domain equals $\{c^{\mathbf{G}} \mid c$ is an individual constant of $\bigcup_n L_n\}$. By the invariance of α for generated subframes, it follows that $TC(\mathbf{G}, \mathrm{w}_1^{\mathbf{G}}, \ldots, \mathrm{w}_n^{\mathbf{G}}) \models [\mathrm{w}_1/x_1, \ldots, \mathrm{w}_n/x_n]\alpha$. Now, define a function f from $TC(\mathbf{G}, \mathrm{w}_1^{\mathbf{G}}, \ldots, \mathrm{w}_n^{\mathbf{G}})$ to the limit $\mathbf{F}$ of the chain $\mathbf{F}_1, \mathbf{F}_2, \ldots$ by setting $f(c^{\mathbf{G}}) =_{\mathrm{def}} c^{\mathbf{F}}$. We claim that f is a p-morphism from $TC(\mathbf{G}, \mathrm{w}_1^{\mathbf{G}}, \ldots, \mathrm{w}_n^{\mathbf{G}})$ onto $TC(\mathbf{F}, \mathrm{w}_1^{\mathbf{F}}, \ldots, \mathrm{w}_n^{\mathbf{F}})$.

That f is *well-defined* follows from the fact that, if $c_1^{\mathbf{G}} = c_2^{\mathbf{G}}$, then, for a suitably large $n \in IN$, $c_1 \in L_n$ and $c_2 \in L_n$ and, moreover, $\mathbf{G}_n \models c_1 = c_2$. Therefore, since $\mathbf{G}_n - RP(L_n) - \mathbf{F}_n$, $\mathbf{F}_n \models c_1 = c_2$ and hence $\mathbf{F} \models c_1 = c_2$.

That f is *onto* follows from the observation that the domain of $TC(\mathbf{F}, \mathrm{w}_1^{\mathbf{F}}, \ldots, \mathrm{w}_n^{\mathbf{F}})$ equals $\{c^{\mathbf{F}} \mid c$ is an individual constant of $\bigcup_n L_n\}$.

That f is a *homomorphism* follows like above. If $\mathbf{G} \models Rc_1c_2$, then, for a suitably large $n \in IN$, $\mathbf{G}_n \models Rc_1c_2$, whence $\mathbf{F}_n \models Rc_1c_2$ (and $\mathbf{F} \models Rc_1c_2$); Rc_1c_2 being restricted positive (as was $c_1 = c_2$).

Finally, if $\mathbf{F} \models Rc_1x[v]$, then $v = c_2^{\mathbf{F}}$ for some individual constant c of $\bigcup_n L_n$ (recall the construction of the languages L_n: some k_{c_1v} will do); and hence — because $\mathbf{G} \models Rc_1k_{c_1v}$ — $k_{c_1v}^{\mathbf{G}}$ is the required R-successor of $c_1^{\mathbf{G}}$ which is mapped by f onto v. This completes the verification that f is a p-morphism.

Finally, since α is preserved under p-morphic images, $TC(\mathbf{F}, \mathrm{w}_1^{\mathbf{F}}, \ldots, \mathrm{w}_n^{\mathbf{F}}) \models [\mathrm{w}_1/x_1, \ldots, \mathrm{w}_n/x_n]\alpha$ and — again by the invariance of α for generated subframes — $\mathbf{F} \models [\mathrm{w}_1/x_1, \ldots, \mathrm{w}_n/x_n]\alpha$. By Tarski's theorem, $\mathbf{F}_1 \models [\mathrm{w}_1/x_1, \ldots, \mathrm{w}_n/x_n]\alpha$ and so, at last, $\mathbf{F}_1^! \models \alpha[w_1, \ldots, w_n]$. QED.

15.5 *Corollary.* Each formula in *P1* is logically equivalent to a restricted positive L_0-formula with the same free variable.

Proof. Any formula in *P1* is invariant for generated subframes and preserved under p-morphic images, because its defining modal formula is (cf. 2.2 and 2.18). QED.

The next series of notions and results concerns $\bar{P}1$.

15.6 *Definition.* An L_0-sentence α is *preserved under generated subframes* if, for all frames $\mathbf{F}_1$, $\mathbf{F}_2$ such that $\mathbf{F}_1 \subseteq_{\rightarrow} \mathbf{F}_2$, $\mathbf{F}_2 \models \alpha$ only if $\mathbf{F}_1 \models \alpha$.

15.7 *Definition.* For any first-order language L containing L_0, the *existentially restricted L-formulas* are the formulas of L constructed from atomic formulas and their negations using $\wedge, \vee$, $\forall$ and restricted existential quantifiers $\exists y(Rty \wedge$, where t is an individual constant or a variable distinct from t.

15.8 *Theorem* (R. I. Goldblatt, S. Feferman [21]). An L_0-sentence is preserved under generated subframes if and only if it is equivalent to an existentially restricted L_0-sentence.

Proof. The essential ideas of this proof are similar to those in the proof of theorem 15.4. Therefore, the exposition will be more sketchy now: only the new considerations will be stressed.

First, if α is an existentially restricted L_0-formula with the free variables $x_1, \ldots, x_n$, then, for all frames $\mathbf{F}_1$ ($= \langle W_1, R_1 \rangle$) and $\mathbf{F}_2$ with $\mathbf{F}_1 \subseteq_{\rightarrow} \mathbf{F}_2$ and for all $w_1, \ldots, w_n \in W_1$, $\mathbf{F}_2 \models \alpha[w_1, \ldots, w_n]$ only if $\mathbf{F}_1 \models \alpha[w_1, \ldots, w_n]$. (Use induction with respect to the construction of α.) Clearly, then, one half of the theorem has been established.

Now, consider any L_0-sentence α which is preserved under generated subframes. Let $ER(\alpha)$ be $\{ \beta \mid \beta$ is an existentially restricted L_0-sentence such that $\alpha \models \beta\}$. It suffices to prove that $ER(\alpha) \models \alpha$, in order to establish the remaining half of the theorem.

Starting with $\mathbf{F}_0$ such that $\mathbf{F}_0 \models ER(\alpha)$(for which it is to be shown that $\mathbf{F}_0 \models \alpha$), two elementary chains $\mathbf{F}_0, \mathbf{F}_1, \mathbf{F}_2, \ldots$ and $\mathbf{G}_0, \mathbf{G}_1, \mathbf{G}_2, \ldots$ are constructed. The only salient points are the construction principle and the final reasoning.

First, each finite subset of $\{\alpha\} \cup \{ \neg\beta \mid \beta$ is an existentially restricted L_0-sentence such that $\mathbf{F}_0 \models \neg \beta\}$ has a model, as is shown by a (by now) familiar argument. Let $\mathbf{G}_0$ be a model for the whole set. This provides the starting point for the construction: see the illustration on the next page, where (i) $\mathbf{F}_0$ and $\mathbf{G}_0$ are L_0-structures, and (ii) $\mathbf{G}_0 - ER(L_0) - \mathbf{F}_0$: i.e., each existentially restricted L_0-sentence which is true in $\mathbf{G}_0$ is also true in $\mathbf{F}_0$.

frames:

$$\begin{tikzcd} \mathbf{F}_0 \\ \mathbf{G}_0 \arrow[u] \end{tikzcd}$$

languages: L_0,

Next, let $\mathbf{F}_n$, $\mathbf{G}_n$ and L_n be given such that $\mathbf{F}_n$ and $\mathbf{G}_n$ are L_n-structures satisfying $\mathbf{G}_n - ER(L_n) - \mathbf{F}_n$. For each individual constant c of L_n and each w in the domain of $\mathbf{G}_n$ such that $\mathbf{G}_n \models Rcx[w]$, add a new individual constant w to L_n, to obtain a language L_n^1. Then expand $\mathbf{G}_n$ to an L_n^1-structure $\mathbf{G}_n^1$ by interpreting each w as w.

Each finite subset of $\Delta = \{ \beta \mid \beta$ is an existentially restricted sentence of L_n^1 such that $\mathbf{G}_n^1 \models \beta\}$ has a model which is an expansion of $\mathbf{F}_n$. To see this, let $\beta_1, \ldots, \beta_k \in \Delta$, containing (in all) $\mathrm{w}_1, \ldots, \mathrm{w}_s$ from $L_n^1 - L_n$. Then, for suitable L_n-constants $c_1, \ldots, c_s$ and for individual variables $x_1, \ldots, x_s$ not occurring in $\beta_1, \ldots, \beta_k$, $\exists x_1(Rc_1x_1 \wedge \ldots \wedge \exists x_s(Rc_sx_s \wedge [x_1/\mathrm{w}_1, \ldots, x_s/\mathrm{w}_s](\beta_1 \wedge \ldots \wedge \beta_k)$ holds in $\mathbf{G}_n$; whence this existentially restricted L_n(!)-sentence holds in $\mathbf{F}_n$. Etcetera. It follows that Δ has a model $\mathbf{F}_n^1$ which is an L_n^1-structure as well as an L_n-elementary extension of $\mathbf{F}_n$, while $\mathbf{G}_n^1 - ER(L_n^1) - \mathbf{F}_n^1$. The situation may be pictured as follows.

frames:

$$\begin{tikzcd} \mathbf{F}_n & \prec & \mathbf{F}_n^1 \\ \mathbf{G}_n, \arrow[u] & \mathbf{G}_n^1 \arrow[ur] & \end{tikzcd}$$

languages: L_n, L_n^1, L_n^1,

Now,to return to the lower chain. For each world w in the domain of $\mathbf{F}_n^1$, add a new individual constant w to L_n^1, to obtain a language L_{n+1}. Expand $\mathbf{F}_n^1$ to an L_{n+1}-structure $\mathbf{F}_{n+1}$. Then each finite subset of $\Gamma = \{\neg\beta \mid \beta$ is an existentially restricted sentence of L_{n+1} such that $\mathbf{F}_{n+1} \models \neg\beta\} \cup \{Rc\mathrm{w} \mid c$ is a constant in L_n^1 and w is a constant in $L_{n+1} - L_n^1$ such that $\mathbf{F}_{n+1} \models Rc\mathrm{w}\}$ has a model which is an expansion of $\mathbf{G}_n^1$.

For $\neg\beta_1, \ldots, \neg\beta_k$ as described and $Rc_1\mathrm{w}_1, \ldots, Rc_s\mathrm{w}_s$, the argument goes as follows. Let $v_1, \ldots, v_t$ be the constants from $L_{n+1} - L^1_n$ occurring in $\neg\beta_1, \ldots, \neg\beta_k$ which are not among $w_1, \ldots, w_s$. Now, assume that $\neg\beta_1 \wedge \ldots \wedge \neg\beta_k \wedge Rc_1w_1 \wedge \ldots \wedge Rc_sw_s$ is not satisfiable in any expansion of $\mathbf{G}^1_n$. Then, for variables $x_1, \ldots, x_s$ and $y_1, \ldots, y_t$ not occurring in this formula,

$$\mathbf{G}^1_n \models \forall y_1 \ldots \forall y_t \forall x_1 \ldots \forall x_s([x_1/w_1, \ldots, x_s/w_s, y_1/v_1, \ldots, y_t/v_t] \\ (\beta_1 \vee \ldots \vee \beta_k \vee \neg Rc_1w_1 \vee \ldots \vee \neg Rc_sw_s).$$

But, since this is an existentially restricted L^1_n-sentence, it would have to hold in $\mathbf{F}^1_n$ as well; which contradicts the assumptions on these formulas. It follows that Γ has a model $\mathbf{G}_{n+1}$ which is an L_{n+1}-structure as well as an L^1_n-elementary extension of $\mathbf{G}^1_n$, while $\mathbf{G}_{n+1} - ER(L_{n+1}) - \mathbf{F}_{n+1}$. Put graphically, this amounts to the following picture:

frames:

$$\begin{tikzcd} \mathbf{F}_n & \prec & \mathbf{F}^1_n, & \mathbf{F}_{n+1} \\ \mathbf{G}_n, \arrow[u] & \mathbf{G}^1_n \arrow[ur] & \prec & \mathbf{G}_{n+1}, \arrow[u] \end{tikzcd}$$

languages: $L_n, \quad L^1_n, \quad L^1_n, \qquad L_{n+1},$

It remains to give the final reasoning. Since $\mathbf{G}_0 \models \alpha$, α holds in the limit $\mathbf{G}$ of the chain $\mathbf{G}_0, \mathbf{G}_1, \mathbf{G}_2, \ldots$. The substructure $\mathbf{G}^*$ of $\mathbf{G}$ whose domain is $\{c^{\mathbf{G}} \mid c$ is a constant of $\bigcup_n L_n\}$ is a generated subframe of $\mathbf{G}$; by the manner in which these constants were chosen in the construction of the structures $\mathbf{F}^1_n$. Then, α holds in $\mathbf{G}^*$, being preserved under generated subframes.

Now define a function f from $\mathbf{G}^*$ to the limit $\mathbf{F}$ of the chain $\mathbf{F}_0, \mathbf{F}_1, \mathbf{F}_2, \ldots$ by putting $f(c^{\mathbf{G}}) =_{\text{def}} c^{\mathbf{F}}$. It is easily checked that f is an isomorphism between $\mathbf{G}^*$ and $\mathbf{F}$. (Use the fact that every (negation of an) atomic sentence of $\bigcup_n L_n$ which is true in $\mathbf{G}$ is also true in $\mathbf{F}$ — by construction —and that each object in $\mathbf{F}$ has a name in $\bigcup_n L_n$ — again by construction.) But, then, $\mathbf{F} \models \alpha$ and hence $\mathbf{F}_0 \models \alpha$. QED.

15.9 *Definition.* An L_0-sentence α is *preserved under p-morphisms* if, for all p-morphisms f from some frame $\mathbf{F}_1$ onto a frame $\mathbf{F}_2$, $\mathbf{F}_1 \models \alpha$ only if $\mathbf{F}_2 \models \alpha$.

15.10 *Definition.* For any first-order language L containing L_0, the *positive L-formulas allowing universal restrictions* (abbreviated, somewhat misleadingly, as: the *universally restricted positive L-formulas*) are the formulas of L constructed from atomic formulas and $\bot$ using $\wedge, \vee, \exists, \forall(!)$, and restricted universal quantifiers $\forall y(Rty \rightarrow$; where t is an individual constant or a variable distinct from y.

The following result may be compared to Lyndon's homomorphism theorem (cf. Chang & Keisler [17]).

15.11 *Theorem.* An L_0-sentence is preserved under p-morphisms if and only if it is equivalent to a universally restricted positive L_0-sentence.

Proof. The argument is similar to the preceding one. The only differences are the following.

(i) L_n-structures $\mathbf{F}_n$, $\mathbf{G}_n$ are given such that each universally restricted positive L_n-sentence which is true in $\mathbf{G}_n$ is also true in $\mathbf{F}_n$. Constants w may be added now *for each w* in the domain of $\mathbf{G}_n$, to obtain L_n^1. (In view of the presence of unrestricted existential quantifiers, there is no need to restrict attention to those w for which $\mathbf{G}_n \models Rcx[w]$ for some L_n-constant c.) One easily shows that each finite set of L_n^1-sentences which are true in the expanded structure $\mathbf{G}_n^1$ is satisfied in some expansion of $\mathbf{F}_n$, etc.

(ii) L_n-structures $\mathbf{F}_n$, $\mathbf{G}_n$ are given as in (i), together with an L_n^1-structure $\mathbf{F}_n^1$ which is an L_n-elementary extension of $\mathbf{F}_n$ such that each universally restricted positive L_n^1-sentence which is true in $\mathbf{G}_n^1$ is also true in $\mathbf{F}_n^1$. One has to add *two* kinds of new individual constants now. First, for each L_n^1-constant c and each w in the domain of $\mathbf{F}_n^1$ such that $\mathbf{F}_n^1 \models Rcx[w]$, a new constant k_{cw} is added. (These serve to ensure that the function to be defined from the chain limit $\mathbf{G}$ to the chain limit $\mathbf{F}$ will satisfy the "backward" clause in the definition of a p-morphism.) Moreover, for each w in the domain of $\mathbf{F}_n^1$, a new constant w is added (in order to ensure that the same function will become *onto*). This yields a language L_{n+1}. Each finite set of *negated* universally restricted positive L_{n+1}-sentences which are true in the expanded structure $\mathbf{F}_{n+1}$, together with a finite set of L_{n+1}-sentences of the form Rck_{cw} (for k_{cw} in $L_{n+1} - L_n^1$) has a model which is an expansion of $\mathbf{G}_n^1$. This is shown by reductio ad absurdum, using the fact that both restricted and unrestricted universal quantifiers are allowed in the construction of universally restricted positive sentences.

(iii) The chain limits $\mathbf{G}$ and $\mathbf{F}$ are given as $\bigcup_n L_n$-structures. Because $\mathbf{G}_0 \models \alpha$, $\mathbf{G} \models \alpha$. The function f from $\mathbf{G}$ to $\mathbf{F}$ defined as in the above proof turns out to be a p-morphism onto (by inspection of the construction). Clearly, then, $\mathbf{F} \models \alpha$ and hence $\mathbf{F}_0 \models \alpha$. QED.

The case of preservation under *disjoint unions* has turned out to be more intractable. It becomes necessary to construct systems of elementary chains simultaneously. Moreover, the syntactic formulation is less elegant. Note that, e.g., $\forall x \exists y Rxy$ is preserved under disjoint unions (as is $\forall x \exists y Ryx$!), whereas $\forall x \forall y Rxy$ is not. In fact, we have no more than the following *conjecture*:

An L_0-sentence is preserved under disjoint unions if and only if it is equivalent to a sentence of the form $\forall x\alpha$, where $\alpha = \alpha(x)$ is constructed from atomic formulas and their negations, $\wedge$, $\vee$, $\exists$ and restricted universal quantifiers of the types $\forall y(Rty \rightarrow$ and $\forall y(Ryt \rightarrow$; where t is an individual constant or a variable distinct from y.

Instead, the following property of modal formulas is treated (cf. corollary 2.15).

15.12 *Definition.* An L_0-sentence α is *invariant for disjoint unions* if, for all sets $\{\mathbf{F}_i \mid i \in I\}$ of frames, $\Sigma\,\{\mathbf{F}_i \mid i \in I\} \models \alpha$ iff, for all $i \in I$, $\mathbf{F}_i \models \alpha$.

15.13 *Definition.* For any first-order language L containing L_0, the *two-way restricted L-formulas* are the formulas of L in which all occurrences of quantifiers are restricted according to the types: $\forall y(Rty \rightarrow$, $\forall y(Ryt \rightarrow$, $\exists y(Rty \wedge$ and $\exists y(Ryt \wedge$; where t is an individual constant or a variable distinct from y.

15.14 *Theorem.* An L_0-sentence is invariant for disjoint unions if and only if it is equivalent to an L_0-sentence of the form $\forall x\alpha$, where $\alpha = \alpha(x)$ is a two-way restricted L_0-formula.

Proof. Let $\{\mathbf{F}_i \mid i \in I\}$ be any set of frames, and let $w_1, \ldots, w_n \in W_j$ for some $j \in I$. For each two-way restricted L_0-formula $\alpha = \alpha(x_1, \ldots, x_n)$, it holds that $\Sigma\,\{\mathbf{F}_i \mid i \in I\} \models \alpha[\langle j, w_1\rangle, \ldots, \langle j, w_n\rangle]$ iff $\mathbf{F}_j \models \alpha[w_1, \ldots, w_n]$. (Use induction on the complexity of α.) The assertion that $\forall x\alpha(x)$ as above is invariant for disjoint unions follows as a direct corollary.

Now, let α be any L_0-sentence which is invariant for disjoint unions. Define $2R(\alpha)$ as $\{\forall x\beta \mid \beta = \beta(x)$ is a two-way restricted L_0-formula such that $\alpha \models \forall x\beta\}$. It will be shown that $2R(\alpha) \models \alpha$. Then, by compactness, for some finite number $\forall x_1\beta_1, ..., \forall x_k\beta_k$ of sentences in $2R(\alpha)$, $\forall x_1\beta_1, ..., \forall x_k\beta_k \models \alpha$. Changing bound variables in some appropriate fashion yields $\forall x\beta'_1, ..., \forall x\beta'_k$ in $2R(\alpha)$ implying α as well. It follows that α is logically equivalent to $\forall x\beta'_1 \wedge ... \wedge \forall x\beta'_k$; i.e., to $\forall x(\beta'_1 \wedge ... \wedge \beta'_k)$, which is a sentence of the required kind.

For a start, let $\mathbf{F}_0^1 \models 2R(\alpha)$. Now $\mathbf{F}_0^1$ — and indeed any frame — may be considered to be (isomorphic to) a disjoint union of its *components*. By this, the minimal subframes of $\mathbf{F}_0^1$ are meant which are closed under R-successors and R-predecessors. In fact, in any frame $\mathbf{F}$ $(= \langle W, R\rangle)$, each $w \in W$ generates a component $\overline{TC}(\mathbf{F}, w)$, which is the subframe of $\mathbf{F}$ whose domain W_w is found inductively as follows. $\bar{S}_0(w) = \{w\}$, $\bar{S}_{n+1}(w) = \bar{S}_n(w) \cup \{v \in W \mid$ for some $u \in \bar{S}_n(w)$, Ruv or $Rvu\}$. (Cf. definition 2.7.) $W_w =_{def} \bigcup_n \bar{S}_n(w)$. Note that each component of $\mathbf{F}$ is a generated subframe of $\mathbf{F}$; but the converse is not true. Also note that, if $v \in W_w$, then $\overline{TC}(\mathbf{F}, v) = \overline{TC}(\mathbf{F}, w)$.

This being understood, write $\mathbf{F}_0^1$ as a disjoint union of its components in some way. Say, $\mathbf{F}_0^1 = \Sigma\{\mathbf{F}_{0w}^1 \mid w \in I\}$, where I is contained in the domain of $\mathbf{F}_0^1$, and $\mathbf{F}_{0w}^1 = \overline{TC}(\mathbf{F}_0^1, w)$. Now, consider any $\mathbf{F}_{0w}^1$. Add a constant w to L_0 to obtain L_w, and expand $\mathbf{F}_{0w}^1$ to an L_w-structure $\mathbf{F}_{0w}$ by interpreting w as w. Carrying out this procedure for all $w \in I$ yields an expansion $\mathbf{F}_0$ of $\mathbf{F}_0^1$ $(\mathbf{F}_0 = \Sigma\{\mathbf{F}_{0w} \mid w \in I\})$ which is an L_w-structure for each $w \in I$. Each finite subset of $\{\alpha\} \cup \{\beta \mid \beta$ is a two-way restricted sentence of L_w such that $\mathbf{F}_0 \models \beta\}$ has a model (the argument showing this must be routine by now), and so the whole set has a model $\mathbf{G}_w$. (By the way, note that there are no two-way restricted L_0-*sentences*; but the constant w makes it possible to have such sentences in L_w.) Define G0 to be the set of all $\mathbf{G}_w (w \in I)$ obtained in this way. For uniformity of notation, write "$L_0(\mathbf{G}_w)$" for the above "L_w". Then the following starting position has been reached.

(i) For each $\mathbf{G} \in$ G0, $\mathbf{G}$ — $2R(L_0(\mathbf{G}))$ — $\mathbf{F}_0$; i.e., each two-way restricted sentence of $L_0(\mathbf{G})$ which is true in $\mathbf{G}$ is also true in $\mathbf{F}_0$,

(ii) For different $\mathbf{G}$ in G0, the languages $L_0(\mathbf{G})$ have disjoint sets of individual constants,

and

(iii) For each $\mathbf{G}$ in $G0$, all constants from $L_0(\mathbf{G})$ are interpreted in one single component of $\mathbf{F}_0$, in which no interpretations occur of constants from different languages $L_0(\mathbf{G}')$.

The general construction starts from the situation as described just now, but for an arbitrary n (instead of 0).

Consider any $\mathbf{G} \in Gn$. For each $L_n(\mathbf{G})$-constant c and each w in the domain of $\mathbf{G}$ such that $\mathbf{G} \models Rcx \vee Rxc[w]$, add a *new* constant w to $L_n(\mathbf{G})$. This yields a language $L_n^1(\mathbf{G})$, and $\mathbf{G}$ may be expanded to an $L_n^1(\mathbf{G})$-structure $\mathbf{G}^1$ by interpreting each w as w. This procedure is to be followed with respect to all $\mathbf{G} \in Gn$; bearing in mind that the new constants w are to be chosen in such a way as to keep the languages $L_n^1(\mathbf{G})$ disjoint. Each finite subset of $\bigcup_{\mathbf{G} \in Gn} \{\beta \mid \beta$ is a two-way restricted sentence of $L_n^1(\mathbf{G})$ which is true in $\mathbf{G}^1\}$ has a model which is an expansion of $\mathbf{F}_n$. (Use the fact that the two-way restricted $L_n(\mathbf{G})$-sentences are closed under restricted existential quantification of two kinds.) Therefore, the whole set has a model $\mathbf{F}_n^1$ satisfying, for each $\mathbf{G} \in Gn$,

(i) $\mathbf{F}_n^1$ is an $L_n^1(\mathbf{G})$-structure,
(ii) $\mathbf{F}_n$ is an $L_n(\mathbf{G})$-elementary substructure of $\mathbf{F}_n^1$,
(iii) $\mathbf{G}_n^1 \; - \; 2R(L_n^1(\mathbf{G})) \; - \; \mathbf{F}_n^1$, and
(iv) all constants of $L_n^1(\mathbf{G})$ are interpreted in one single component of $\mathbf{F}_n^1$ (viz. the one where those from $L_n(\mathbf{G})$ were interpreted).

Now, for the other direction, consider again any $\mathbf{G} \in Gn$. Take new constants w for all those elements w in the domain of $\mathbf{F}_n^1$ such that $\mathbf{F}_n^1 \models Rcx \vee Rxc[w]$ for some constant c of $L_n^1(\mathbf{G})$ to obtain $L_n^2(\mathbf{G})$. Moreover, for each component of $\mathbf{F}_n^1$ in which no interpretation of a constant occurs as yet, choose any element w in that component and add a *new* individual constant w to L_0; to obtain new languages L_w. Expand $\mathbf{F}_n^1$ to a structure $\mathbf{F}_{n+1}$ for all these extended languages by interpreting each new w as the corresponding w. Repeat the procedure of the beginning of this proof with respect to the last-mentioned L_w's. This yields structures $\mathbf{G}_w$ as before. Then, for each of the $L_n^2(\mathbf{G})$'s, consider the set of two-way restricted $L_n^2(\mathbf{G})$-sentences which are true in $\mathbf{F}_{n+1}$. As above, each finite subset of this set has a model which is an expansion of $\mathbf{G}^1$. (Use the fact that the $L_n^2(\mathbf{G})$-sentences are closed under negations). Therefore, the whole set has a model $\mathbf{G}^2$ satisfying

(i) $\mathbf{G}^2$ is an $L_n^2(\mathbf{G})$-structure,
(ii) $\mathbf{G}^1$ is an $L_n^1(\mathbf{G})$-elementary substructure of $\mathbf{G}^2$, and
(iii) $\mathbf{G}^2 - 2R(L_n^2(\mathbf{G})) - \mathbf{F}_{n+1}$.

Rename $L_n^2(\mathbf{G})$ as "$L_{n+1}(\mathbf{G})$", and define $G(n+1)$ as the collection of all structures $\mathbf{G}^2(\hat{\mathbf{G}} \in Gn)$ and $\mathbf{G}_w$ as obtained above. Thus, the situation indexed by n, has reappeared with index $n+1$.

This construction yields elementary chains, each of which starts from some $\mathbf{G}$ in some Gn; with chain limits $C(\mathbf{G})$. From the (sets of formulas used in the) construction, it follows immediately that $C(\mathbf{G})$ is a $\bigcup_n L_n(\mathbf{G})$-structure in which the interpretations of the constants in $\bigcup_n L_n(\mathbf{G})$ form a component $C'(\mathbf{G})$.

Moreover, the disjoint union $\mathbf{G}^*$ of these components is isomorphic to the limit $\mathbf{F}$ of the chain $\mathbf{F}_0$, $\mathbf{F}_1$, $\mathbf{F}_2$, ... (The obvious isomorphism is defined as in previous proofs.) Now, by construction, $\mathbf{G} \models \alpha$ for each $\mathbf{G}$ in each Gn, and, therefore, $C(\mathbf{G}) \models \alpha$. By the invariance of α for disjoint unions, $C'(\mathbf{G}) \models \alpha$; and, for the same reason, $\mathbf{G}^* \models \alpha$. It follows that $\mathbf{F} \models \alpha$ and hence $\mathbf{F}_0^1 \models \alpha$. QED.

In spite of the failure to prove the expected preservation result for disjoint unions, a combined result has been found which suffices for the intended application to $\bar{P}1$. (Cf. theorem 15.4.)

15.15 *Theorem.* An L_0-sentence is preserved under generated subframes, disjoint unions and p-morphisms if and only if it is equivalent to an L_0-sentence of the form $\forall x\, \beta$, where $\beta = \beta(x)$ is a restricted positive L_0-formula.

Proof. The direction from left to right follows immediately from previous observations. For the converse, suppose that the L_0-sentence α is preserved under generated subframes, disjoint unions and p-morphisms. It will be shown that $\bar{R}\bar{P}(\alpha) \models \alpha$; where $\bar{R}\bar{P}(\alpha) = \{\forall x\, \beta \mid \beta = \beta(x)$ is a restricted positive L_0-formula such that $\alpha \models \forall x\, \beta\}$. As in the proof of theorem 15.14, this will yield the required equivalent for α.

Consider any $\mathbf{F}_0^1$ such that $\mathbf{F}_0^1 \models \bar{R}\bar{P}(\alpha)$. Take constants w for each w in the domain of $\mathbf{F}_0^1$; which yields languages $L_w = L_0 \cup \{\mathrm{w}\}$. Expand $\mathbf{F}_0^1$ to an $\bigcup_w L_w$-structure $\mathbf{F}_0$ by interpreting each w as w. Now, for any w, each finite subset of $\Gamma_w = \{\alpha\} \cup \{\neg\beta \mid \beta$ is a restricted positive sentence of L_w

such that $\mathbf{F}_0 \models \neg\beta\}$ has a model. Otherwise, e.g., $\{\alpha, \neg\beta_1, ..., \neg\beta_k\}$ has no model. In other words, $\alpha \models \beta_1 \vee ... \vee \beta_k$ and, therefore, $\alpha \models \forall x[x/\mathrm{w}](\beta_1 \vee ... \vee \beta_k)$. But, then, $\mathbf{F}_0^1 \models \forall x[x/\mathrm{w}](\beta_1 \vee ... \vee \beta_k)$ (this sentence being in $\bar{R}\bar{P}(\alpha)$); contradicting the fact that $\mathbf{F}_0^1 \models [x/\mathrm{w}](\neg\beta_1 \wedge ... \wedge \neg\beta_k)[w]$. It follows that Γ_w has a model $\mathbf{G}_w$. Defining $L_0(\mathbf{G}_w)$ to be L_w and $\mathbf{G}0$ to be the set of all $\mathbf{G}_w$'s obtained in this way yields the following point of departure.

For each $\mathbf{G} \in \mathbf{G}0$,

(i) $\mathbf{F}_0$ is an $L_0(\mathbf{G})$-structure,

(ii) $\mathbf{G} - RP(L_0(\mathbf{G})) - \mathbf{F}_0$, (where the relation "$- RP(L) -$" was defined in the proof of theorem 15.4), and

(iii) for different structures $\mathbf{G}$ in $\mathbf{G}0$, the languages $L_0(\mathbf{G})$ have disjoint sets of individual constants.

Again, elementary chains will be constructed, according to the following principle.

Let $\mathbf{G}n$, $\mathbf{F}_n$ and, for each $\mathbf{G} \in \mathbf{G}n$, $L_n(\mathbf{G})$ be given such that the above three clauses hold with "n" substituted for "0". Consider any $\mathbf{G} \in \mathbf{G}n$. For each w in the domain of $\mathbf{G}$ such that $\mathbf{G} \models Rcx[w]$ for some $L_n(\mathbf{G})$-constant c, add a *new* constant w to $L_n(\mathbf{G})$ to obtain a language $L_n^1(\mathbf{G})$. $\mathbf{G}$ may then be expanded to an $L_n^1(\mathbf{G})$-structure $\mathbf{G}^1$ by interpreting each w as w. This procedure is repeated for each $\mathbf{G} \in \mathbf{G}n$; taking care to keep the sets of individual constants of different languages $L_n^1(\mathbf{G})$ disjoint.

Now, let $\Delta_n(\mathbf{G})$ be the set of restricted positive $L_n^1(\mathbf{G})$-sentences which are true in $\mathbf{G}^1$ ($\mathbf{G} \in \mathbf{G}n$). Each finite subset of $\Delta_n = \bigcup_{\mathbf{G} \in \mathbf{G}n} \Delta_n(\mathbf{G})$ has a model which is an expansion of $\mathbf{F}_n$. For, consider $\beta_1^1, ..., \beta_{k_1}^1 \in \Delta_n(\mathbf{G}_1)$, ..., $\beta_1^m, ..., \beta_{k_m}^m \in \Delta_n(\mathbf{G}_m)$. Let $\beta_1^1, ..., \beta_{k_1}^1$ contain the constants $\mathrm{w}_1^1, ..., \mathrm{w}_{s_1}^1$ from $L_n^1(\mathbf{G}_1) - L_n(\mathbf{G}_1)$ and ... and $\beta_1^m, ..., \beta_{k_m}^m$ the constants $\mathrm{w}_1^m, ..., \mathrm{w}_{s_m}^m$ from $L_n^1(\mathbf{G}_m) - L_n(\mathbf{G}_m)$. Note that no constant w_j^i can be equal to a constant w_l^k (where $i \neq k$); because the languages $L_n^1(\mathbf{G}_1)$,..., $L_n^1(\mathbf{G}_m)$ have disjoint sets of individual constants. Therefore, it suffices to show that each separate set $\beta_1^i, ..., \beta_{k_i}^i$ has a model which is an expansion of $\mathbf{F}_n$. To see this, let $c_1, ..., c_{s_i}$ be constants of $L_n(\mathbf{G}_i)$ such that $\mathbf{G}_i \models Rc_jx[w_j^i]$ $(1 \leqslant j \leqslant s_i)$. For suitable new variables $x_1, ..., x_s$, it follows that

$$\mathbf{G}_i \models \exists x_1(Rc_1x_1 \wedge ... \wedge \exists x_{s_i}(Rc_{s_i}x_{s_i} \wedge [x_1/\mathrm{w}_1^i, ..., x_{s_i}/\mathrm{w}_{s_i}^i](\beta_1^i \wedge ... \wedge \beta_{k_i}^i).$$

This is a restricted positive $L_n(\mathbf{G}_i)$-sentence; which holds in $\mathbf{F}_n$ by clause (ii) above. Finally, then, Δ_n has a model $\mathbf{F}_n^1$ satisfying, for each $\mathbf{G} \in \mathbf{G}n$,

(i) $\mathbf{F}_n^1$ is an $L_n^1(\mathbf{G})$-structure,
(ii) $\mathbf{F}_n$ is an $L_n(\mathbf{G})$-elementary substructure of $\mathbf{F}_n^1$, and
(iii) $\mathbf{G}^1 - RP(L_n^1(\mathbf{G})) - \mathbf{F}_n$.

Now for the other direction. Consider any language $L_n^1(\mathbf{G})(\mathbf{G} \in \mathbf{G}n)$. For each individual constant c of $L_n^1(\mathbf{G})$ and each w in the domain of $\mathbf{F}_n^1$ such that $\mathbf{F}_n^1 \models Rcx[w]$, add a *new* constant k_{cw} to $L_n^1(\mathbf{G})$; to obtain $L_n^2(\mathbf{G})$. (Note that different $L_n^2(\mathbf{G})$'s retain disjoint sets of individual constants.) Expand $\mathbf{F}_n^1$ to an $L_n^2(\mathbf{G})$-structure (for each $\mathbf{G}$) $\mathbf{F}_n^2$ by interpreting each constant k_{cw} as w. Each finite subset of $\Sigma_n(\mathbf{G}) = \{\neg\beta \mid \beta$ is a restricted positive sentence of $L_n^2(\mathbf{G})$ such that $\mathbf{F}_n^2 \models \neg\beta\} \cup \{Rck_{cw} \mid k_{cw} \in L_n^2(\mathbf{G}) - L_n^1(\mathbf{G})\}$ has a model which is an expansion of $\mathbf{G}^1$. (Use previous arguments, exploiting the fact that the restricted positive $L_n^1(\mathbf{G})$-sentences are closed under restricted universal quantification.) Therefore, $\Sigma_n(\mathbf{G})$ has a model $\mathbf{G}^2$, which is an $L_n^2(\mathbf{G})$-structure with $\mathbf{G}^1$ as an $L_n^1(\mathbf{G})$-elementary substructure and, moreover, $\mathbf{G}^2 - RP(L_n^2(\mathbf{G})) - \mathbf{F}_n^2$.

Next, take new individual constants $\mathbf{w}$ for all those elements w in the domain of $\mathbf{F}_n^2$ which have not been named yet by any constant in any $L_n^2(\mathbf{G})$. Expand $\mathbf{F}_n^2$ to a structure $\mathbf{F}_{n+1}$ by interpreting each such $\mathbf{w}$ as w. Since, by the construction up to now, the original frame $\mathbf{F}_0^1$ is an L_0-elementary substructure of $\mathbf{F}_{n+1}$, it holds that $\mathbf{F}_{n+1} \models \overline{RP}(\alpha)$. But, then, the procedure followed in the construction of G0 may be repeated with respect to $\mathbf{F}_{n+1}$ and these new constants $\mathbf{w}$, to obtain models $\mathbf{G}_w$ for α with corresponding languages $L_{n+1}(\mathbf{G}_w)(=L_0 \cup \{\mathbf{w}\})$ such that $\mathbf{G}_w - RP(L_{n+1}(\mathbf{G}_w)) - \mathbf{F}_{n+1}$.

Defining $\mathbf{G}(n+1)$ to be $\{\mathbf{G}^2 \mid \mathbf{G} \in \mathbf{G}n\} \cup \{\mathbf{G}_w \mid \mathbf{G}_w$ is constructed as indicated in the preceding paragraph$\}$ and renaming $L_n^2(\mathbf{G})$ as "$L_{n+1}(\mathbf{G}^2)$" yields the situation from which we started, but now with $n + 1$ instead of n. (The three relevant clauses are obviously satisfied.)

The described construction creates a set of elementary chains, each beginning with a structure $\mathbf{G}$ in some $\mathbf{G}n$; as well as a single elementary chain $\mathbf{F}_0, \mathbf{F}_1, \mathbf{F}_2, \ldots$. Since $\mathbf{G} \models \alpha$ for each $\mathbf{G}$ in each $\mathbf{G}n$, α will be true in each chain limit $C(\mathbf{G})$ of a chain starting from $\mathbf{G}$.

As in previous proofs, an obvious function $f_{\mathbf{G}}$ may be defined as follows. For any constant c of $\bigcup_n L_n(\mathbf{G})$, set $f_{\mathbf{G}}(c^{C(\mathbf{G})}) =_{\text{def}} c^{\mathbf{F}}$; where $\mathbf{F}$ is the limit of the chain $\mathbf{F}_0, \mathbf{F}_1, \mathbf{F}_2, \ldots$. A routine check reveals that $f_{\mathbf{G}}$ is a p-morphism from the generated subframe of $C(\mathbf{G})$ whose domain is $\{c^{C(\mathbf{G})} \mid c \in \bigcup_n L_n(\mathbf{G})\}$ onto some generated subframe of $\mathbf{F}$. The union f of all these

p-morphisms $f_{\mathbf{G}}$ is itself a p-morphism from a generated subframe of the disjoint union of all limits $C(\mathbf{G})$ onto $\mathbf{F}$. (That it is *onto* follows from the continuous creation of new languages L_w during the construction.)

Now, α holds in all limits $C(\mathbf{G})$. Therefore, it holds in the disjoint union of these limits; being preserved under such unions. Moreover, being preserved under generated subframes, it holds in the generated subframe of this disjoint union which is the domain of f. Applying its third preservation property (involving p-morphisms) shows that α holds in $\mathbf{F}$. Finally, by the fundamental theorem on elementary chains, $\mathbf{F}_0^1 \models \alpha$. QED.

15.16 *Corollary.* Each sentence in $\bar{P}1$ is logically equivalent to an L_0-sentence of the form $\forall x\, \alpha$, where $\alpha = \alpha(x)$ is a restricted positive L_0-formula.

The next step on the road towards a complete syntactic characterization of $\bar{P}1$ seems to be the formulation of an analogous preservation result for *ultrafilter extensions* (cf. corollary 14.8). No such result has been found yet: only some partial answers will be presented below. In fact, it should be borne in mind that the class of L_0-sentences which are preserved under ultrafilter extensions need not be recursively enumerable: the search could be futile from the start.

First, recall the relevant facts from chapter 2 (2.24 - 2.27). Note especially that $\mathbf{F}$ is isomorphically embedded in $ue(\mathbf{F})$ through the mapping $w \mapsto \{X \subseteq W \mid w \in X\}$ $(= w^*)$. Thus, for all practical purposes, $\mathbf{F}$ may be considered to be a subframe of $ue(\mathbf{F})$. This implies that existential L_0-sentences are preserved under ultrafilter extensions; but such a non-specific result is hardly exciting, of course. A little more information is provided by corollary 15.19 below.

15.17 *Definition.* Let u be any fixed individual variable. An *$r(u)$-formula* is any L_0-formula obtained from atomic formulas of the forms Rux, Rxu, $u = x$, $x = u$ (where x is distinct from u) and atomic formulas in which u does not occur, by applying Boolean operators and two types of restricted existential quantification, taking formulas $\alpha(u)$ to

$$\exists y(Ruy \wedge \alpha(y)) \quad \text{or} \quad \exists y(Ryu \wedge \alpha(y));$$

where y does not occur in $\alpha(u)$.

E.g., for $i \in IN$ and for any variable x distinct from u, the formula $R^i ux$ is an $r(u)$-formula (cf. chapter 9).

15.18 *Lemma.* For any frame $\mathbf{F}$ $(= \langle W, R\rangle)$ with $w_1, ..., w_k \in W$, any $r(u)$-formula $\alpha = \alpha(u, x_1, ..., x_k)$, and any ultrafilter U on $\mathbf{F}$,

$ue(\mathbf{F}) \models \alpha[U, w_1^*, ..., w_k^*]$ if and only if $\{v \in W \mid \mathbf{F} \models \alpha[v, w_1, ..., w_k]\} \in U$.

Proof. This equivalence follows by induction on the construction of α.

α is Rx_ix_j: $ue(\mathbf{F}) \models Rx_ix_j[U, w_1^*, ..., w_k^*]$ iff $R_{\mathbf{F}}w_i^*w_j^*$ iff Rw_iw_j iff $\mathbf{F} \models Rx_ix_j[v, w_1, ..., w_k]$ (for any $v \in W$) iff $\{v \in W \mid \mathbf{F} \models \alpha[v, w_1, ..., w_k]\} \in U$.

α is $x_i = x_j$: similarly.

α is Rux_i: $ue(\mathbf{F}) \models Rux_i[U, w_1^*, ..., w_k^*]$ iff $R_{\mathbf{F}}\, Uw_i$ iff (by an easy calculation) $\{v \in W \mid Rvw_i\} \in U$ iff $\{v \in W \mid \mathbf{F} \models Rux_i\, [v, w_1, ..., w_k]\} \in U$.

α is Rx_iu: similarly; using the fact that $R_{\mathbf{F}}w_iU$ iff $\{v \in W \mid Rw_iv\} \in U$.

α is $u = x_i$: $ue(\mathbf{F}) \models u = x_i[U, w_1^*, ..., w_k^*]$ iff $U = w_i^*$ iff $\{w_i\} \in U$ iff $\{v \in W \mid \mathbf{F} \models u = x_i[v, w_1, ..., w_k]\} \in U$.

α is $x_i = u$: similarly.

α is $\neg\beta$ or $\beta \wedge \gamma$: these cases follow by standard arguments, using the characteristic properties of ultrafilters.

α is $\exists y(Ruy \wedge \beta(y))$: $ue(\mathbf{F}) \models \exists y(Ruy \wedge \beta(y))[U, w_1^*, ..., w_k^*]$ iff, for some ultrafilter V in $W_{\mathbf{F}}$ such that $R_{\mathbf{F}}UV$, $ue(\mathbf{F}) \models \beta[V, w_1^*, ..., w_k^*]$ iff (by the induction hypothesis), for some V in $W_{\mathbf{F}}$ such that $R_{\mathbf{F}}UV$, $\{v \in W \mid \mathbf{F} \models \beta[v, w_1, ..., w_k]\} \in V$.

Now apply the following general principle:

Let $\alpha = \alpha(y, y_1, ..., y_k)$ be an L_0-formula. For any $w_1, ..., w_k \in W$ and any $U \in W_{\mathbf{F}}$, $\{v \in W \mid \exists z \in W(Rvz \;\&\; \mathbf{F} \models \alpha[z, w_1, ..., w_k])\} \in U$ if and only if, for some $V \in W_{\mathbf{F}}$ such that $R_{\mathbf{F}}UV$, $\{v \in W \mid \mathbf{F} \models \alpha[v, w_1, ..., w_k]\} \in V$.

The standard deduction establishing this principle is omitted here. One merely has to use the Ultrafilter Extension Theorem with respect to a suitable filter of subsets of W.

Then, the list of equivalences may be continued with: iff $\{v \in W \mid \exists z \in W(Rvz \;\&\; \mathbf{F} \models \beta[z, w_1, ..., w_k])\} \in U$; i.e., iff $\{v \in W \mid \mathbf{F} \models \exists y(Ruy \wedge \beta(y))[v, w_1, ..., w_k]\} \in U$.

α is $\exists y(Ryu \wedge \beta(y))$: this is proved analogously; but now using the dual principle: $\{v \in W \mid \exists z \in W(Rzv \ \&\ \mathbf{F} \models \alpha[z, w_1, ..., w_k])\} \in U$ if and only if, for some $V \in W_{\mathbf{F}}$ such that $R_{\mathbf{F}}VU$, $\{v \in W \mid \mathbf{F} \models \alpha[v, w_1, ..., w_k]\} \in V$. **QED**.

15.19 *Corollary*. For any $r(u)$-formula $\alpha = \alpha(u, x_1, ..., x_k)$ and for any frame $\mathbf{F}$ $(= \langle W, R\rangle)$ with $w, w_1, ..., w_k \in W$, $ue(\mathbf{F}) \models \alpha[w^*, w_1^*, ..., w_k^*]$ if and only if $\mathbf{F} \models \alpha[w, w_1, ..., w_k]$.

Proof. By Lemma 15.18, $ue(\mathbf{F}) \models \alpha[w^*, w_1^*, ..., w_k^*]$ iff $\{v \in W \mid \mathbf{F} \models \alpha[v, w_1, ..., w_k]\} \in w^*$, and this is the case iff $w \in \{v \in W \mid \mathbf{F} \models \alpha[v, w_1, ..., w_k]\}$ iff $\mathbf{F} \models \alpha[w, w_1, ..., w_k]$. **QED**.

This corollary implies that any L_0-sentence obtained from an $r(u)$-formula by existential quantification is preserved under ultrafilter extensions. This extends the above observation about the preservation of existential L_0-sentences. Unfortunately, it does not exhaust the class of sentences wich are preserved under ultrafilter extensions, however. E.g., both $\forall xRxx$ and $\forall x\forall yRxy$ have this property too (although $\forall x \neg Rxx$ does not). That $\forall x\neg Rxx$ is not preserved may be seen from the example following corollary 2.26. For, $\langle IN, <\rangle \models \forall x\neg Rxx$, but $ue(\langle IN, <\rangle) \models \forall x \exists y(Rxy \wedge Ryy)$. On the other hand, if $\mathbf{F} \models \forall xRxx$, then $ue(\mathbf{F}) \models \forall xRxx$. For, let $U \in W_{\mathbf{F}}$ and let X be any set in U. $\{w \in W \mid \exists v \in X\ Rwv\}$ contains X and is, therefore, in U. In other words, $R_{\mathbf{F}}UU$ holds.

Pending further research in this direction, the syntactic structure of $\bar{P}1$ remains opaque.

AN APPENDIX (TENSE LOGIC)

Some of the notions introduced in this chapter belong to *tense logic* rather than *modal logic*. By "tense logic" the propositional language is meant with added operators G ("it will always be the case that") and F ("it will be the case that"), as well as H ("it has always been the case that") and P ("it was the case that"). Semantically, frames $\mathbf{F} = \langle W, R\rangle$ now function as *time structures*; with W as the set of *moments* and R as the ordering relation "*earlier than*". In the truth definition, G and F are treated like L and M, respectively; while H and P become dual notions using the *converse* relation of R. (For a fuller exposition, cf. Prior [64].) The modal theory of the present work generalizes to tense logic in a rather uneventful manner. Only minor modifications are needed. One example is that tense-logical formulas need not be preserved under generated subframes. They are preserved under disjoint unions, however; and in fact they are *invariant* for disjoint unions (cf. definition 15.12). It was no coincidence, then, that *two-way* restricted L_0-formulas turned up in the statement of theorem 15.14. Similarly, tense-logical formulas are not always preserved under p-morphisms, but they are preserved under $\bar{p}$-*morphisms*; i.e., p-morphisms satisfying the additional condition (cf. definition 2.16) (iii) "for all $w \in W_1$ and $v \in W_2$, $R_2vf(w)$ only if there exists a $u \in W_1$ such that R_1uw and $f(u) = v$." Clearly, an obvious variant of theorem 15.11 may be formulated and proven. Finally, ultrafilter extensions are acceptable in tense logic as they are. Note, e.g., the characteristic fact that, in definition 15.17, quantifiers $\exists y(Ryu \wedge$ played the same role as quantifiers $\exists y(Ruy \wedge$. Apart from the philosophical virtue that tense logic may be the system of intensional logic for which the Kripke semantics has most natural plausibility, there is also the technical virtue that tense logic often provides easy examples of phenomena (e.g., incompleteness: cf. Thomason [79]) which are much harder to find in the more restricted language of modal logic. A full exposition of the varieties of tense-logical semantics, including their Correspondence Theory, may be found in the present author's book *The Logic of Time*, Reidel, Dordrecht, 1982, Synthese Library, vol. 156.

CHAPTER XVI

MODALLY DEFINABLE CLASSES OF FRAMES

In chapter 14 it was asked which elementary classes of frames are modally definable. The more general question with which this chapter will be concerned is "When is a class of frames modally definable at all?". As it turns out, the most elegant answer obtained here is not one to that particular question, but to the more special question which classes of frames are modally definable by means of a *canonical* set of modal formulas (cf. definition 6.8).

To start with the general question; the answer is surely not obvious, since we are in fact asking a question of second-order logic: when is a class of frames definable by means of universal second-order sentences (of an admittedly rather simple kind)? It is natural to turn to the "nearest" first-order language, so to speak; which is a *two-sorted* predicate logic extending L_0 (whose variables range over "individuals") with variables for "sets of individuals" (cf. Enderton [20]). General frames will be structures for this language whose first domain is W and whose second domain is $\mathbf{W}$. Then the well-known characterization of first-order definable classes of structures (due to H. J. Keisler) could, hopefully, be applied in the present case. This is in fact the road which will be followed in chapters 17 and 18; but, for the special case of modal formulas, a more elegant approach is available. It consists in an *algebraization* of the subject, after which Birkhoff's equally well-known characterization of algebraic varieties may be applied (cf. Grätzer [32]):

The smallest equationally definable class of algebras to contain a class $\mathbf{K}$ of algebras is $\mathbf{HSP}(\mathbf{K})$, where

H(X) is the class of *homomorphic images* of members of X,
S(X) is the class of *subalgebras* of members of X, and
P(X) is the class of *direct products* of members of X.

The duality underlying this approach has been sketched in chapter 4. It is described in full detail in Goldblatt [30].

As a first indication of the utility of this approach, a result will be formulated ([30], theorem 12.11) about modal definability of classes of *general* frames. (Cf. theorem 16.1 below.) Recall the duality developed in chapter 4.

16.1 *Theorem* (R. I. Goldblatt). A class **K** of general frames is modally definable if and only if it is closed under generated subframes, disjoint unions, p-morphic images and Stone Representations, while its complement is closed under Stone Representations. If, in addition, **K**'s complement is closed under ultraproducts, then **K** is definable by means of a single modal formula (and conversely).

Proof. If **K** is modally definable, then the above closure conditions hold by the results of chapter 4. If a single modal formula defines **K**, then a routine argument based upon the generalization of the theorem of Łoś to these ultraproducts establishes that **K**'s complement is closed under ultraproducts.

Conversely, let **K** satisfy the closure conditions of the first statement. It will be shown that **K** consists of exactly the general frames in which Th_{mod} (**K**) (i.e., $\{\varphi \mid \varphi$ is a modal formula true in all members of **K**$\}$) holds. One direction of this is trivial. For the other, suppose that $\langle \mathbf{F}, W\rangle \models \Sigma = Th_{mod}$ (**K**). It is to be proven that $\langle \mathbf{F}, W\rangle$ actually belongs to **K**.

On the algebraic side, $\langle \mathbf{F}, W\rangle^+ \models \Sigma^+$; where Σ^+ is the set of polynomial identities $\bar{\varphi} = 1$ (cf. the definition of "polynomial transcription" in chapter 4) corresponding to the modal formulas φ in Σ. Obviously, Σ^+ is a set of polynomial identities which are true in all algebras in $\mathbf{K}^+ = \{\langle \mathbf{F}, W\rangle^+ \mid \langle \mathbf{F}, W\rangle \in \mathbf{K}\}$. In fact, up to a trivial difference, Σ^+ *is* the algebraic theory of $\mathbf{K}^+$. For, if the polynomial identity $t_1 = t_2$ holds in all algebras of $\mathbf{K}^+$, then an obvious corresponding modal formula $\varphi_1 \leftrightarrow \varphi_2$ belongs to Σ.

Now, it is easy to see that, for any class X of algebras, the smallest equational variety to contain X is the class of all algebras in which the algebraic theory of X holds. By Birkhoff's theorem, this variety also equals

$\mathrm{HSP}(X)$. Applying this to $\langle \mathbf{F}, W\rangle^+$ and $\mathbf{K}^+$, one obtains that $\langle \mathbf{F}, W\rangle^+$ must belong to $\mathrm{HSP}(\mathbf{K}^+)$. In other words, there exist general frames $\langle \mathbf{F}_i, W_i\rangle^+$ $(i \in I)$ in $\mathbf{K}$ such that $\langle \mathbf{F}, W\rangle^+$ is a homomorphic image of some subalgebra $\mathbf{A}$ of $\prod_{i \in I} \langle \mathbf{F}_i, W_i\rangle^+$. By lemma 4.10, $\prod \langle \mathbf{F}_i, W_i\rangle^+$ is isomorphic to $(\Sigma\{\langle \mathbf{F}_i, W_i\rangle \mid i \in I\})^+$; where the disjoint union $\langle \mathbf{F}_u, W_u\rangle = \Sigma\{\langle \mathbf{F}_i, W_i\rangle \mid i \in I\}$ belongs to $\mathbf{K}$ ($\mathbf{K}$ being closed under disjoint unions). It follows that $\mathbf{A}$ is isomorphic to some subalgebra $\mathbf{A}'$ of $\langle \mathbf{F}_u, W_u\rangle^+$ and hence that $\langle \mathbf{F}, W\rangle^+$ is a homomorphic image of $\mathbf{A}'$. In a diagram we have

$$\langle \mathbf{F}_u, W_u\rangle^+ \quad \supseteq \quad \mathbf{A}' \longrightarrow \langle \mathbf{F}, W\rangle^+.$$

Using lemma 4.9, this may be extended to

$$\begin{array}{ccccc} \langle \mathbf{F}_u, W_u\rangle^+ & \supseteq & \mathbf{A}' & \longrightarrow & \langle \mathbf{F}, W\rangle^+ \\ | & & | & & \\ SR(\langle \mathbf{F}_u, W_u\rangle^+) & \longrightarrow & SR(\mathbf{A}'), & & \end{array}$$

where $SR(\mathbf{A}')$ is a p-morphic image of $SR(\langle \mathbf{F}_u, W_u\rangle^+)$. Using lemma 4.8, the second diagram may be completed to

$$\begin{array}{ccccc} \langle \mathbf{F}_u, W_u\rangle^+ & \supseteq & \mathbf{A}' & \longrightarrow & \langle \mathbf{F}, W\rangle^+ \\ | & & | & & | \\ SR(\langle \mathbf{F}_u, W_u\rangle^+) & \longrightarrow & SR(\mathbf{A}') & \longleftarrow & SR(\langle \mathbf{F}, W\rangle^+), \end{array}$$

where $SR(\mathbf{F}, W\rangle^+)$ is isomorphic to a generated subframe of $SR(\mathbf{A}')$. Now, $SR(\langle \mathbf{F}_u, W_u\rangle^+)$ is in $\mathbf{K}$ ($\mathbf{K}$ being closed under Stone Representations); whence $SR(\mathbf{A}')$ is in $\mathbf{K}$ ($\mathbf{K}$ being closed under p-morphic images) and also $SR(\langle \mathbf{F}, W\rangle^+) \in \mathbf{K}$ (because $\mathbf{K}$ is closed under generated subframes and isomorphic images: a special case of p-morphic images). Finally, then, $\langle \mathbf{F}, W\rangle$ must belong to $\mathbf{K}$, since $\mathbf{K}$'s complement is closed under Stone Representations.

It remains to consider the second statement of the theorem. It suffices to prove that, if $\mathbf{K}$ is modally definable and $\mathbf{K}$'s complement is closed under ultraproducts, then a single modal formula defines $\mathbf{K}$. A consideration similar to one in the proof of lemma 14.1 shows that it then suffices to prove that the smallest equational variety to contain $\mathbf{K}^+$ (i.e., the class of algebras in which Σ^+ holds; which properly contains $\mathbf{K}^+$, because it has non-set algebras among its members) is defined by a finite set of polynomial identities. For that purpose, it is to be shown that the complement of that variety is closed under ultraproducts (after which Keisler's theorem is applicable as in the proof of 14.1). Now, let $\mathbf{A}_i$ ($i \in I$) be modal algebras in which Σ^+ does not hold, and let U be an ultrafilter on I. It is to be shown that $\Pi_U \mathbf{A}_i \not\models \Sigma^+$. But this follows easily from the observation that each $SR(\mathbf{A}_i)$ belongs to the complement of $\mathbf{K}$, whence $\Pi_U SR(\mathbf{A}_i)$ belongs to that class as well, and — again like in the proof of 14.1 — $\Pi_U \mathbf{A}_i \cong \Pi_U SR(\mathbf{A}_i)^+ \cong (\Pi_U SR(\mathbf{A}_i))^+$. QED.

In their paper [31], R. I. Goldblatt and S. K. Thomason used this kind of consideration to characterize the modally definable classes of *frames*. Since the concepts (and the result) they ended up with are typically "proof-generated", the story had best be told starting upside down, so to speak.

Let $\mathbf{K}$ be a class of frames. Clearly, if $\mathbf{K}$ is modally definable at all, then it will be defined by $Th_{mod}(\mathbf{K})$. Consider, then, any frame $\mathbf{F}$ such that $\mathbf{F} \models Th_{mod}(\mathbf{K})$. What closure conditions on $\mathbf{K}$ are required to ensure that $\mathbf{F} \in \mathbf{K}$? The obvious procedure is to change to the corresponding algebras, as in the previous proof. Set $\Sigma = Th_{mod}(\mathbf{K})$. Then $\mathbf{F}^+(= \langle \mathbf{F}, P(W)\rangle^+)$ $\models \Sigma^+$ and, as before, $\mathbf{F}^+ \in \mathrm{HSP}(\mathbf{K}^+)$. Following the previous argument still further, a disjoint union $\mathbf{F}_u$ of frames in $\mathbf{K}$ is found such that $\mathbf{F}^+$ is a homomorphic image of some subalgebra $\mathbf{A}'$ of $\mathbf{F}_u^+$.

First requirement: $\mathbf{F}_u$ is to belong to $\mathbf{K}$; i.e., $\mathbf{K}$ is to be *closed under disjoint unions*.

Then we had $SR(\mathbf{A}')$ as a p-morphic image of $SR(\mathbf{F}_u^+)$; i.e., of $ue(\mathbf{F}_u)$. But, this is of no help, since a requirement to the effect that $\mathbf{K}$ be closed under *ultrafilter extensions* (and, of course, *p-morphic images*) is too strong: not all modally definable classes $\mathbf{K}$ satisfy it; as we have seen in chapter 2. Still, this fact is something to be noted (cf. theorem 16.4).

To bridge the gap between $\mathbf{F}$ and $\mathbf{F}_u$, for which one only knows that $\mathbf{F}^+$ is a homomorphic image of some subalgebra $\mathbf{A}'$ of $\mathbf{F}_u^+$ — or, in terms of the above duality, that $SR(\mathbf{F}^+)$ is isomorphic to a generated subframe of $SR(\mathbf{A}')$ — Goldblatt and Thomason introduced the following concept.

16.2 *Definition.* A frame $\mathbf{F}$ is *SA-based* on a frame $\mathbf{F}_2$ if, for some W such that $\langle \mathbf{F}_2, W\rangle$ is a general frame, $\mathbf{F}_1$ is a subframe of the underlying frame of $SR(\langle \mathbf{F}_2, W\rangle^+)$ satisfying the following three conditions, for all U, $V \in W_1$,

(i) R_1UV iff, for all $X \in W$, if $l(X) \in U$, then $X \in V$,
(ii) for all $X \in W$, if $l(X) \notin U$, then, for some $U' \in W_1$, R_1UU' and $X \notin U'$, and
(iii) for all $Y \subseteq W_1$, there exists an $X \in W$ such that $Y = W_1 \cap \{Z \mid Z \text{ is an ultrafilter on } \langle \mathbf{F}, W\rangle^+ \text{ containing } X\}$.

Note that, in terms of the above situation, $\mathbf{A}'$, being a subalgebra of $\mathbf{F}_u^+$, equals $\langle \mathbf{F}_u, W\rangle^+$ for some W. Moreover, $\mathbf{F}$ is isomorphically embedded in $ue(\mathbf{F})$, which is itself isomorphic to a generated subframe of $SR(\langle \mathbf{F}_u, W\rangle^+)$. A straightforward calculation shows that $\mathbf{F}$ is isomorphic to a frame which is SA-based on $\mathbf{F}_u$. In fact, the authors prove the following general assertion.

16.3 *Lemma.* For any frame $\mathbf{F}$ and any class $\mathbf{K}$ of frames, $\mathbf{F}^+ \in HS(\mathbf{K}^+)$ if and only if $\mathbf{F}$ is isomorphic to a frame which is SA-based on some frame in $\mathbf{K}$.

An obvious consequence of this lemma is

16.4 *Theorem* (R. I. Goldblatt & S. K. Thomason). A class $\mathbf{K}$ of frames is modally definable if and only if it is closed under disjoint unions, frames which are SA-based upon some frame in $\mathbf{K}$, and isomorphic images.

Another way to exploit the above argument is as follows. Recall the *canonical* sets of definition 6.11. Such sets are preserved in passing from a *descriptive* general frame in which they hold to the underlying frame. Canonical modal definability can be characterized using much less *ad hoc* notions:

16.5 *Theorem.* A class $\mathbf{K}$ of frames is definable by means of a canonical set of modal formulas if and only if it is closed under generated subframes, disjoint unions, p-morphic images and ultrafilter extensions, while its complement is closed under ultrafilter extensions as well.

Proof. If $\mathbf{K}$ is canonically definable, then it is closed under ultrafilter extensions (the other closure properties follow already from its being modally definable). For, let the canonical set Σ define $\mathbf{K}$. Consider any $\mathbf{F} \in \mathbf{K}$; i.e., $\mathbf{F} \models \Sigma$. By the results of chapter 4, $SR(\mathbf{F}^+) = \langle ue(\mathbf{F}), W\rangle \models \Sigma$. But this general frame is descriptive and hence $ue(\mathbf{F}) \models \Sigma$, because Σ is canonical.

If $\mathbf{K}$ satisfies the mentioned closure properties, then the above argument may be used once more. For any $\mathbf{F}$ such that $\mathbf{F} \models Th_{mod}(\mathbf{K})$, it was shown that $SR(\mathbf{F}^+)$ is isomorphic to a generated subframe of some p-morphic image of some general frame $SR(\mathbf{F}_u^+)$ with $\mathbf{F}_u \in \mathbf{K}$. Then one follows the diagram: $\mathbf{F}_u \in \mathbf{K}$, whence $ue(\mathbf{F}_u) \in \mathbf{K}$. Now, $SR(\mathbf{F}_u^+) = \langle ue(\mathbf{F}_u), W\rangle$ for some W; so, for its p-morphic images $\langle \mathbf{F}', W'\rangle$, $\mathbf{F}' \in \mathbf{K}$. For any generated subframe $\langle \mathbf{F}'', W''\rangle$ of $\langle \mathbf{F}', W'\rangle$, $\mathbf{F}'' \in \mathbf{K}$ and the same holds for the $\mathbf{F}'''$ in its isomorphic copies $\langle \mathbf{F}''', W'''\rangle$. In particular, it follows that, since $SR(\mathbf{F}^+) = \langle ue(\mathbf{F}), W'''\rangle$ is such a copy, $ue(\mathbf{F}) \in \mathbf{K}$. Finally, $\mathbf{F} \in \mathbf{K}$; because $\mathbf{K}$'s complement is closed under ultrafilter extensions. QED.

As an easy consequence we have the result by R. I. Goldblatt and S. K. Thomason which was given already as theorem 14.7:

16.6 *Corollary.* If a class $\mathbf{K}$ of frames is closed under (L_0-) elementary equivalence, then $\mathbf{K}$ is modally definable if and only if $\mathbf{K}$ is closed under generated subframes, disjoint unions and p-morphic images, while its complement ${}^C\mathbf{K}$ is closed under ultrafilter extensions.

16.7 *Corollary* (K. Fine). If a set Σ of modal formulas is complete and preserved under elementary equivalence, then it is canonical.

Proof. $FR(\Sigma)$ is closed under p-morphic images and elementary equivalence. Therefore, by theorem 8.9, it is closed under ultrafilter extensions. Thus, theorem 16.5 becomes applicable: $FR(\Sigma) = FR(\Delta)$ for some canonical set Δ of modal formulas. Since Σ is complete (cf. definition 6.4) and since, moreover, $\Sigma \models_f \delta$ for all $\delta \in \Delta$, it holds that $\Sigma \models_{gf} \delta$ for all $\delta \in \Delta$ (or, equivalently, $\Delta \subseteq ML(\Sigma)$). Now, let $\langle \mathbf{F}, W\rangle$ be any descriptive general frame in which Σ holds. Clearly, then, $\langle \mathbf{F}, W\rangle \models \Delta$; whence $\mathbf{F} \models \Delta$, Δ being canonical. It follows that $\mathbf{F} \in FR(\Sigma)$ as well. QED.

It was shown in the proof of 16.5 that any canonical set of modal formulas is preserved under ultrafilter extensions. The converse does not hold. Every modal formula in $\bar{M}1$ is preserved under ultrafilter extensions, by virtue of theorem 8.9. But not every modal formula in $\bar{M}1$ is canonical: some of them are not even complete (cf. chapter 6). What we do have are the following two results.

16.8 *Lemma.* Σ is canonical if and only if it is preserved in going from a general frame to the underlying frame of its Stone Representation.

Proof. If Σ is canonical and $\langle \mathbf{F}, W\rangle \models \Sigma$, then $SR((\langle \mathbf{F}, W\rangle^{+}) \models \Sigma$. Moreover, $SR(\langle \mathbf{F}, W\rangle^{+})$ is descriptive and so Σ holds on its underlying frame.

If Σ has the described preservation property, then, if $\langle \mathbf{F}, W\rangle \models \Sigma$, where $\langle \mathbf{F}, W\rangle$ is descriptive, Σ holds in the underlying frame of $SR(\langle \mathbf{F}, W\rangle^{+})$; which is isomorphic to $\mathbf{F}$ itself. QED.

16.9 *Lemma.* Σ is preserved under ultrafilter extensions if and only if $C(\Sigma)$ is canonical.

Proof. From left to right, this follows from theorem 16.5. From right to left, it follows from the fact that canonical sets are preserved under ultrafilter extensions. QED.

No results have been formulated for classes of frames definable by means of a *single* modal formula. One would like to know if the obvious additional requirement that the complementary class be closed under ultraproducts is sufficient. It is clearly necessary.

APPENDIX

HIGHER-ORDER CORRESPONDENCE

Modal formulas define second-order (Π^1_1) conditions upon the alternative relation in all cases, and first-order conditions in some. In the perspective of abstract model theory, two possible generalizations arise here.

Instead of the first-order target language, one may consider suitable extensions. For instance, in various earlier-mentioned cases of non-first-order definability, the relational condition expressed was definable in $L_{\omega_1\omega}$: first-order logic with *countable* conjunctions and disjunctions. Not all modal formulas become definable, however, in this wider reals. E.g., Löb's Axiom defined a form of well-foundedness, a property which is known to be beyond $L_{\omega_1\omega}$, or indeed any language of the $L_{\infty\omega}$-family. On the other hand, well-foundedness is itself definable in "weak second-order logic" L^2, allowing quantification over *finite* sets of individuals. Thus, various wider classes of definability could be considered for modal formulas, short of the full Π^1_1. And in fact, even the latter general case itself invites further study. For instance, which Π^1_1-sentences admit of modal definitions?

Given the general lack of semantic characterizations for such higher logics, such characterizations for their modal fragments are also difficult to obtain. One relevant observation might be that both $L_{\omega_1\omega}$ and L^2 have the property of *invariance for partial isomorphism*, in the sense of abstract model theory. It will be of interest to study this preservation condition on modal formulas. Indeed, no counter-examples have been discovered yet; but these do exist in tense logic. (The rationals and the reals form a classical example of partially isomorphic order structures. Still, there exists a formula of tense logic expressing Dedekind Completeness, which is valid in the latter, though not in the former frame.)

On the other side, the modal propositional language could itself be strengthened, notably by the introduction of *propositional quantifiers* $\forall p$, $\exists p$; which have occurred at various places in the literature. Thus, e.g., $\forall p(LMp \rightarrow \exists q\, MLq)$ would become an admissible formula; but also e.g., $L \exists p\, Mp \rightarrow M \forall q\, MLq$. Actually, there is a choice here, whether to allow the propositional quantifiers within the scope of modal operators, or not. We consider the latter, more restricted option henceforth.

In the usual manner, a prenex hierarchy arises here, with all propositional quantifiers in front; of which the original modal formulas form the Π^1_1-part (universal prefix). The next simplest cases are Σ^1_1 (existential prefix) and Δ^1_2. In fact, the latter case has a reasonable motivation through the "modal rules" proposed by D. Gabbay. For instance, he has observed that irreflexivity, though modally undefinable, is expressed by the following rule on Kripke frames:

"if $\mathbf{F} \models (Lp \wedge \neg p) \rightarrow \varphi[w]$ (with φ p-free), then $\mathbf{F} \models \varphi[w]$".

The general pattern here is that of "$\mathbf{F} \models \varphi[w]$ only if $\mathbf{F} \models \psi[w]$", i.e., an implication of two Π^1_1-formulas: which is Δ^1_2. (It may be written in either of the forms $\forall\exists$, $\exists\forall$.)

Actually, the specific example above is already Σ^1_1, as it amounts to $\forall pq(Lp \wedge \neg p \rightarrow q) \rightarrow \forall q\, q$, i.e., $\forall p(Lp \wedge \neg p \rightarrow \forall q\; q) \rightarrow \forall qq$, i.e., $\forall p(Lp \wedge \neg p \rightarrow \bot) \rightarrow \bot$, i.e. $\exists p(Lp \wedge \neg p)$. Another relevant observation in this area is that implications of the above form $\forall \rightarrow \forall$, if first-order definable at all, already have a first-order definable consequent. We do not go into these specific matters here, but will note a more general issue.

As always in higher-order logic, we are interested in *hierarchy* results. For instance, how much power of first-order definability is added at each stage? It is evident that Σ^1_1-definability adds essentially just all negations of the (local) principles in *P1*, while Δ^1_2 adds conjunctions and disjunctions across *P1* and the latter's "mirror image".

Query: Does the prenex hierarchy of higher-order modal formulas induce an ascending corresponding hierarchy of modally definable principles about the alternative relation?

At least, this possibly ascending hierarchy cannot exhaust all first-order principles; as higher-order modal formulas do retain one basic preservation property from the preceding: their local truth is invariant under passing to generated subframes. (The Generation Theorem yields this consequence all the way up, not just for the original modal Π^1_1-formulas.)

But then, we know what this semantic constraint means in syntactic terms for first-order formulas — thanks to our preservation theorems. These will be the "almost-restricted" ones, consisting of one universal quantifier followed by a compound of atomic formulas with negation, conjunction and restricted quantifiers $\exists y(Rxy \wedge$.

The other preservation properties that were basic for our original modal formulas are lost, however. As was observed earlier, irreflexivity ($\forall x \neg Rxx$) becomes definable, and hence preservation under p-morphisms now fails. Anti-preservation under ultrafilter extensions fails too, because our earlier example $\forall x \exists y(Rxy \wedge Ryy)$ now becomes definable as well. (One straightforward definition uses a propositional quantifier within a modal scope: $M\forall p(Lp \rightarrow p)$. But there is a non-embedded substitute, in the form of $\exists p(Mp \wedge \forall q\ L(p \rightarrow (Lq \rightarrow q)))$.)

Thus, we arrive at two final questions.

Question: Can every almost-restricted first-order formula $\forall x\varphi(x)$ be defined at some level in the modal propositional quantifier hierarchy?

Question: Does the addition of propositional quantifiers within modal scopes add any power of expression?

PART IV

Higher-order Definability

CHAPTER XVII

UNIVERSAL SECOND-ORDER SENTENCES

The language of modal logic can be viewed as a short-hand notation for certain universal second-order sentences. This is the point of the standard translation ST of chapter 3 connecting modal formulas φ (with the proposition letters $p_1, \ldots, p_n$) with universal second-order sentences $\forall x \forall P_1 \ldots \forall P_n ST(\varphi)$. Here $ST(\varphi)$ contains the unary predicate variables $P_1, \ldots, P_n$ as well as a binary predicate constant R.

In this chapter, a second-order language will be considered with that same predicate constant R, identity and predicate variables with an arbitrary number of argument places, allowing for quantification over predicates. Note that *frames* are semantic structures for this language. In fact, most of the results obtained would go through in the presence of just any set of first-order parameters (instead of only R); but the notation remains simpler in the present approach. (Moreover, the spirit of this work is that one studies the — first-order, modal or second-order — theory of a binary relation.) Of this second-order language, only the universal sentences of the form

$$\forall X_1 \ldots \forall X_m \ \varphi,$$

where φ is a first-order sentence in $X_1, \ldots, X_m$, R and $=$, will be at the focus of attention right now; being the most direct generalization of the modal formulas. (Note that L_0-sentences are included as an extreme case.) Such universal sentences are called $\Pi_1^1(R)$-*sentences*, in view of an obvious hierarchy. (Cf. Chang & Keisler [17].) E.g., $\Sigma_1^1(R)$-sentences will be of the form

$$\exists X_1 \ \ldots \ \exists X_m \ \varphi,$$

with φ as above; and $\Pi_2^1(R)$-sentences are of the form

$$\forall X_1 \ldots \forall X_m \exists Y_1 \ldots \exists Y_n\ \varphi;$$

etc. It will be investigated how certain major themes of the preceding parts fare in this wider field.

To begin with, in corollary 2.9 it was shown that M and $\rightarrow$ suffice as primitives when writing modal formulas. Here, the most natural set of primitives is $\{\forall, \rightarrow\}$, however. To see this, first write φ using only $\forall$, $\rightarrow$ and $\neg$. Then eliminate negations in φ by replacing subformulas of the form $\neg\alpha$ by $\alpha \rightarrow \bot$. Now $\bot$ may be eliminated as well by noting that, for any first-order formula φ in $\forall$, $\rightarrow$ and $\bot$, and any unary predicate variable Q not occurring in φ, the following sentence is universally valid:

$$\varphi \leftrightarrow \forall Q([\forall x Qx/\bot]\ \varphi \ \vee\ \forall x Qx).$$

Moreover, the disjunction may be replaced by a combination of implications using the tautology

$$(\alpha \vee \beta) \leftrightarrow ((\alpha \rightarrow \beta) \rightarrow \beta).$$

Another simplification requires a little more thought.

17.1 *Lemma.* Any $\Pi^1_1(R)$-sentence is logically equivalent to a sentence of the form

$$\forall X_1 \ldots \forall X_n \exists y_1 \ldots \exists y_m \forall z_1 \ldots \forall z_k\ \varphi;$$

where φ is a quantifier-free first-order sentence.

Proof. One way to prove this is by using Skolem functions, which are replaced by suitable predicates afterwards. This is the method used in the original report [7]. Here, a trick is applied which is found in J. Barwise: *Admissible Sets and Structures*, (Springer, Berlin, 1975). (Cf. also lemma 11.10.) Consider any $\Pi^1_1(R)$-sentence $\forall X_1 \ldots \forall X_n\ \varphi$. Write φ using $\neg$, $\rightarrow$ and $\forall$. Now, list all subformulas of φ as $\varphi_1, \ldots, \varphi_t$. For any φ_i, define

$A_i = \varphi_i$ if φ_i is atomic,
$A_i = Pz_1 \ldots z_k$, for some *new* k-ary predicate variable P, if φ_i is a compound formula with the free variables $z_1, \ldots, z_k$.

For $\varphi_i = \varphi$ itself, A_i will become a "nullary" predicate variable P_φ.

Form the conjunction of all "connecting sentences" as follows. If $\varphi_j = \neg\varphi_i$, put in $\forall z_1 \ldots \forall z_k(A_j \leftrightarrow \neg A_i)$;where $z_1, \ldots, z_k$ are the relevant free variables. If $\varphi_j = \varphi_i \to \varphi_k$, then put in $\forall z_1 \ldots \forall z_k(A_j \leftrightarrow (A_i \to A_k))$. Finally, if $\varphi_j = \forall z\ \varphi_i$, then put in both $\forall z_1 \ldots \forall z_k(\forall z A_i \to A_j)$ and $\forall z_1 \ldots \forall z_k(A_j \to \forall z A_i)$. Call the resulting conjunction $A(\varphi)$.

Note that each conjunct in $A(\varphi)$ is either equivalent to a universal first-order formula or to one of the form "$\forall z_1 \ldots \forall z_k \exists z$- quantifier-free part". Therefore, $A(\varphi)$ itself is equivalent to a sentence of the form $\forall y_1 \ldots \forall y_m \exists z_1 \ldots \exists z_k\ \psi$, with ψ quantifier-free. Thus, $A(\varphi) \to P_\varphi$ is equivalent to a sentence of the form $\exists y_1 \ldots \exists y_m \forall z_1 \ldots \forall z_k\ \psi$, with ψ quantifier-free.

To arrive at the statement of the lemma, it suffices to note that, for the predicate variables $P_1, \ldots, P_s$ (other than $R, X_1, \ldots, X_n$) occurring in $A_1, \ldots, A_t$, $\forall X_1 \ldots \forall X_n \varphi$ is logically equivalent to $\forall X_1 \ldots \forall X_n \forall P_1 \ldots \forall P_s(A(\varphi) \to P_\varphi)$. (The point is that, whenever $A(\varphi)$ holds, the sentences $\forall z_1 \ldots \forall z_k(A_i \leftrightarrow \varphi_i)$ are true $(1 \leqslant i \leqslant t)$.) QED.

As in the case of modal logic, there is room for a generalized notion of "semantic structure". The original frames are *standard models* for this second-order language (cf. Henkin [33]); but one also needs structures in which the second-order quantifiers range over just a part of all predicates on the domain. Such "general models" will be called *generalized frames*; to distinguish them from the "general frames" of chapter 4. One could define a generalized frame to be any couple $\langle \mathbf{F}, \mathcal{W} \rangle$, where $\mathbf{F}$ is a frame and $\mathcal{W}$ some set of predicates on W. But this notion is rather too liberal, in that one wants certain properties to remain universally valid which would become falsifiable in such structures. One example is the *Comprehension Principle* for first-order formulas $\varphi = \varphi(x_1, \ldots, x_k)$:

$$\exists Y \forall x_1 \ldots \forall x_k(\varphi \leftrightarrow Y x_1 \ldots x_k).$$

Perhaps more illustrative in the present context is the principle of *Universal Instantiation* (to which the Comprehension Principle is intimately related):

$$\forall Y \varphi(Y) \to \varphi(\psi);$$

where Y is a k-ary predicate variable, and ψ a first-order formula with k free variables which is substituted for Y in φ — always taking suitable precautions to avoid confusion of bound variables. This motivates the requirement that $\mathcal{W}$ be closed under "definable predicates". Formally, if $\varphi = \varphi(x_1, \ldots, x_k, y_1, \ldots, y_m, Y_1, \ldots, Y_n)$ is any first-order formula with free individual variables $x_1, \ldots, x_k, y_1, \ldots, y_m$ and (free) predicate variables $Y_1, \ldots,$

Y_n, and if $Z_1, \ldots, Z_n$ are predicates in $\mathbf{W}$ such that Z_i is l-ary if Y_i is $(1 \leqslant i \leqslant n)$ and if, finally, $w_1, \ldots, w_m$ are elements of W, then the predicate $\{\langle v_1, \ldots, v_k\rangle \mid \mathsf{F} \models \varphi\ [v_1, \ldots, v_k, w_1, \ldots, w_m, Z_1, \ldots, Z_n]\}$ is to be in $\mathbf{W}$. This requirement may also be stated inductively:

17.2 *Definition.* A *generalized frame* is an ordered couple $\langle \mathsf{F}, \mathbf{W}\rangle$, where F is a frame $\langle W, R\rangle$ and $\mathbf{W}$ is a set of predicates on W satisfying

(i) R is in $\mathbf{W}$, and so are $=$ and all singletons $\{w\}$ $(w \in W)$
(ii) if an n-ary Z is in $\mathbf{W}$, then $W^n - Z$ is in $\mathbf{W}$ (*complement*)
(iii) if an n-ary Z_1 and an m-ary Z_2 are in $\mathbf{W}$, then the $n + m$-ary predicate $Z_1 \cdot Z_2$ is in $\mathbf{W}$, which is defined as the set of those sequences $\langle w_1, \ldots, w_n, w_{n+1}, \ldots, w_{n+m}\rangle$ in W^{n+m} for which $\langle w_1, \ldots, w_n\rangle \in Z_1$ and $\langle w_{n+1}, \ldots, w_{n+m}\rangle \in Z_2$ (*conjunction*)
(iv) if an $n + 1$-ary Z is in $\mathbf{W}$, then the predicate $\exists Z$ is in $\mathbf{W}$, which is defined as the set of those sequences $\langle w_1, \ldots, w_n\rangle$ in W^n for which a $w \in W$ exists such that $\langle w_1, \ldots, w_n, w\rangle \in Z$ (*projection*)
(v) if an n-ary Z is in $\mathbf{W}$, and $i, j \leqslant n$, then the predicate $\pi_{ij}Z$ is in $\mathbf{W}$, which is defined as the set of those sequences in W^n which, after permuting their i-th and j-th members belong to Z (*permutation*)
(vi) if an $n + 1$-ary Z is in $\mathbf{W}$, then the n-ary predicate IZ is in $\mathbf{W}$, which is defined as the set of those sequences $\langle w_1, \ldots, w_n\rangle$ in W^n for which $\langle w_1, \ldots, w_n, w_n\rangle$ belongs to Z (*identification*).

The closure conditions on $\mathbf{W}$ could be kept more simple for general frames (cf. definition 4.2), because the modal language is so simple: only complement, conjunction and "restricted projection" are needed; involving only *unary* predicates (i.e., subsets of the domain).

Note that the smallest generalized frame $\langle \mathsf{F}, \mathbf{W}\rangle$ on any given frame F has a set $\mathbf{W}$ consisting of exactly those predicates on W which are L_0-definable with parameters in W (cf. definition 9.14). This notion of generalized frame is not the only possible one in the present context —but it will serve as a useful vehicle for the following study.

Universal validity of $\Pi^1_1(R)$-sentences yields nothing new as compared to first-order logic. For, $\forall X_1 \ldots \forall X_n \varphi$ holds in all frames iff it holds in all generalized frames iff the first-order sentence φ is universally valid. Semantic consequence is more interesting, however.

17.3 *Definition.* For any set Σ of $\Pi^1_1(R)$-sentences and any $\Pi^1_1(R)$-sentence φ,

$\Sigma \models \varphi$, if, for all frames $\mathbf{F}$, $\mathbf{F} \models \Sigma$ only if $\mathbf{F} \models \varphi$

$\Sigma \models_g \varphi$ if, for all generalized frames $\langle \mathbf{F}, \mathbf{W} \rangle$, $\langle \mathbf{F}, \mathbf{W} \rangle \models \Sigma$ only if $\langle \mathbf{F}, \mathbf{W} \rangle \models \varphi$.

The notion "$\models$" is highly complex; as was noted in chapter 1. E.g., second-order Zermelo-Fraenkel set theory (ZF^2) may be axiomatized by means of a single $\Pi^1_1(R)$-sentence. Now, since its first-order part is not arithmetically definable (arithmetical truth would be arithmetically definable if this were the case; as was shown in chapter 1), "$\models$" cannot be arithmetically definable either (as a relation between Gödel numbers of $\Pi^1_1(R)$-sentences).

"$\models_g$" turns out to be axiomatizable, however, even for our full second-order language. One merely translates some complete axiom system for first-order logic to second-order logic; e.g., the one given in Enderton [20]. Here the following axioms are taken to fix our thoughts on the matter.

(i) $\forall X(\varphi \rightarrow \psi) \rightarrow (\forall X\varphi \rightarrow \forall X\psi)$

(ii) $\varphi \rightarrow \forall X\varphi$,

provided that X does not occur free in φ;

(iii) $\forall X\varphi \rightarrow [\psi/X]\varphi$,

where X is any k-ary predicate variable and ψ is any first-order formula containing R and possibly predicate variables, whose free individual variables are $y_1, \ldots, y_k$. $[\psi/X]\varphi$ results from φ by replacing subformulas of the form $Xz_1 \ldots z_k$ by $[z_1/y_1, \ldots, z_k/y_k]\psi$; taking bound alphabetic variants whenever confusion of bound variables threatens.

The remaining axioms are the usual first-order ones. The only rule of inference is Modus Ponens. $\Sigma \vdash \varphi$ is then defined in the usual way.

Note that various familiar principles are not derivable in this system. E.g., the Axiom of Choice — needed to handle Skolem normal forms — is not a theorem. It could be added, of course, in forms such as

$$\forall x\exists yRxy \rightarrow \exists Z(\forall x\exists! yZxy \wedge \forall x\ \forall y(Zxy \rightarrow Rxy)).$$

The completeness result will be stated without proof.

17.4 *Theorem.* For all Σ and φ, $\Sigma \vdash \varphi$ if and only if $\Sigma \models_g \varphi$.

For the theory of a binary relation, there are usually three approaches. First, it may be axiomatized by means of a single L_0-sentence. An example is the theory of the rational numbers with the ordering "smaller than". Then, because certain properties of the relation turn out to be undefinable in L_0 (e.g., Dedekind Completeness, or Well-foundedness), there is the second-order approach; often via $\Pi^1_1(R)$-sentences. Examples are second-order Zermelo Fraenkel set theory, or the second-order theory of the natural numbers or the real numbers with the relation "smaller than". Thirdly, there is the first-order approach using axiom *schemata*. It is intermediate, in a sense, between the first two; inspired by the wish to have as much of the second approach as possible without losing the desirable properties of the first. *ZF* itself is an example; and so are first-order Peano Arithmetic and Tarski's "Elementary Geometry". This third intermediate approach can be slightly generalized by allowing predicate variables in the language; as was done above.

The most important notion in chapter 6 was that of *completeness*. Complete modal logics are insensitive to the difference between the "real" $\models$ and the more artificial $\models_g$, so to speak.

17.5 *Definition.* Σ is *complete* if, for all $\Pi^1_1(R)$-sentences φ, $\Sigma \models \varphi$ if and only if $\Sigma \vdash \varphi$.

If Σ is complete, then the set of its L_0-consequences is recursively axiomatizable. It follows, then, that e.g., ZF^2 is not complete. There are also modal examples of such incompleteness phenomena; witness theorem 6.6. The L_0-sentence axiomatizing the theory of the rational numbers with "smaller than" also axiomatizes their $\Pi^1_1(R)$-theory: this follows from its being $\aleph_0$-categorical — and it is, therefore, complete.

The well-known modal completeness theorems (cf. Segerberg [67]) lead to many obvious questions in this area. E.g., $L(Lp \rightarrow p) \rightarrow Lp$ is complete with respect to the class of transitive frames whose converse relation is well-founded. Is, then, the $\Pi^1_1(R)$-sentence defining well-foundedness complete? A similar question concerns the $\Pi^1_1(R)$-sentence defining $\langle IN, <\rangle$ up to isomorphism.

The answer to these general questions is negative. E.g., it is not hard to find a first-order sentence $\varphi = \varphi(R, \in)$ such that, for each model $\langle D,$

$R^*, \in^*\rangle$ of φ in which R^* is well-founded, $\langle D, \in^*\rangle$ is isomorphic to $\langle V_\omega, \in\rangle$ (i.e., the hereditarily finite sets). But, then, for some suitable translation τ of arithmetical sentences α into set-theoretic ones, i.e., into first-order sentences $\tau(\alpha)$ in $\in$ and $=$, the following equivalence will hold. α will be a true statement of arithmetic if and only if "well-foundedness" implies the $\Pi^1_1(R)$-sentence $\forall\in(\varphi(R, \in) \to \tau(\alpha))$. It follows that the set of semantic $\Pi^1_1(R)$-consequences of well-foundedness is not arithmetically definable; by Tarski's Theorem. Still the *monadic* $\Pi^1_1(R)$-theory of, e.g., $\langle IN, <\rangle$ could be complete in the above sense. (The reduction given here uses the binary predicate variable $\in$ essentially.) Indeed, arithmetical complexity can be no obstacle here — as this monadic theory is even *decidable*, by Rabin's Theorem. (This conjecture has been proved in the meantime by H. C. Doets. A related result turns out to hold for the real numbers.)

Next, there is also a parallel to $\bar{M}1$.

17.6 *Definition.* *USO1* is the set of $\Pi^1_1(R)$-sentences φ for which an L_0-sentence ψ exists such that $\varphi \leftrightarrow \psi$ is true on all frames.

$\Sigma^1_1(R)$-sentences are preserved under ultraproducts (cf. Chang & Keisler [17]). By Keisler's Theorem, a class **K** of frames is L_0-elementary iff both **K** and its complement are closed under ultraproducts and isomorphic images. This yields an obvious semantic characterization of *USO1*.

17.7 *Theorem.* A $\Pi^1_1(R)$-sentence is in *USO1* if and only if it is preserved under ultraproducts.

Now, recall lemma 3.12 and its corollary 3.13: any $\Pi^1_1(R)$-sentence of the form $\forall X_1 \ldots \forall X_n \forall y_1 \ldots \forall y_m \varphi$, where φ is constructed from L_0-formulas and atomic formulas of the form $X_i z_1 \ldots z_{n_i}$ $(1 \leqslant i \leqslant n)$ using Boolean operators only, is preserved under ultraproducts.

A related result is the following.

17.8 *Lemma.* Any $\Pi^1_1(R)$-sentence of the form $\forall X_1 \ldots \forall X_n \exists y_1 \ldots \exists y_m \varphi$, where φ is constructed as above, is preserved under ultrapowers.

Proof. The following argument establishes that this kind of sentence is even preserved under L_0-elementary equivalence. Let $\mathbf{F}_1$, $\mathbf{F}_2$ be L_0-

elementarily equivalent frames such that $\mathbf{F}_1 \not\models \forall X_1 \dots \forall X_n \exists y_1 \dots \exists y_m \varphi$. It will be shown that $\mathbf{F}_2 \not\models \forall X_1 \dots \forall X_n \exists y_1 \dots \exists y_m \varphi$. For each $w \in W_2$, add a name $\mathbf{w}$ to L_0; to obtain a language L_0'.

The set $Th_{L_0'}(\mathbf{F}_2) \cup \{\forall y_1 \dots \forall y_m \neg\varphi\}$ is finitely satisfiable. For, consider any finite set of sentences $\psi_1, \dots, \psi_k$ from $Th_{L_0'}(\mathbf{F}_2)$. Suppose that $\{\psi_1, \dots, \psi_k\} \cup \{\forall y_1 \dots \forall y_m \neg\varphi\}$ has no model; i.e., that $\psi_1 \wedge \dots \wedge \psi_k \models \exists y_1 \dots \exists y_m \varphi$. Let $\psi = \psi_1 \wedge \dots \wedge \psi_k$ contain the new constants $\mathbf{w}_1, \dots, \mathbf{w}_s$. Then $\psi' = \exists z_1 \dots \exists z_s[z_1/\mathbf{w}_1, \dots, z_s/\mathbf{w}_s]\psi \models \exists y_1 \dots \exists y_m \varphi$. ψ' is an L_0-sentence, in which none of $X_1, \dots, X_n$ occur. Therefore, it follows that even $\psi' \models \forall X_1 \dots \forall X_n \exists y_1 \dots \exists y_m \varphi$. Since ψ' is true on $\mathbf{F}_2$, it also holds on $\mathbf{F}_1$; these two structures being L_0-elementarily equivalent. But, then, $\mathbf{F}_1 \models \forall X_1 \dots \forall X_n \exists y_1 \dots \exists y_m \varphi$; contrary to the assumption.

Being finitely satisfiable, the whole set $Th_{L_0'}(\mathbf{F}_2) \cup \{\forall y_1 \dots \forall y_m \neg\varphi\}$ has a model $\mathbf{F}_3$ (by the compactness theorem). $\mathbf{F}_2$ is L_0-elementarily embedded in $\mathbf{F}_3$; and may in fact be taken to be an L_0-elementary substructure of $\mathbf{F}_3$. A simple induction with respect to the complexity of formulas shows that, for any formula $\alpha = \alpha(x_1, \dots, x_n)$ constructed from L_0-formulas and atomic formulas of the form $X_i z_1 \dots z_{n_i}$ using Boolean operators, the following holds, for all $w_1, \dots, w_n \in W_2$,

$$\mathbf{F}_3 \models \alpha[w_1, \dots, w_n] \quad \text{iff} \quad \mathbf{F}_2 \models \alpha[w_1, \dots, w_n].$$

An immediate consequence is that $\forall y_1 \dots \forall y_m \neg\varphi$ (where $\neg\varphi$ is a formula α as described), which is true in $\mathbf{F}_3$, is also true in the restriction of $\mathbf{F}_3$ to $\mathbf{F}_2$. This restriction by itself is an expansion of $\mathbf{F}_2$ to a model for $\forall y_1 \dots \forall y_m \neg\varphi$. In other words, $\mathbf{F}_2 \models \exists X_1 \dots \exists X_n \forall y_1 \dots \forall y_m \neg\varphi$. QED.

Lemma 17.8 is the best possible result. For, the following $\Pi_1^1(R)$-sentence is equivalent to one of this form; but it is not preserved under ultra*products*. Let the L_0-sentence α express that R is an irreflexive linear ordering with a first and a last element such that any element except the first has an immediate predecessor and any element except the last has an immediate successor. The sentence

$$\forall P(\alpha \wedge (\forall x \forall y((Px \wedge \neg Py) \to Rxy) \to (\exists x(\neg \exists y Ryx \wedge \neg Px) \vee \exists x(\neg \exists y Rxy \wedge Px) \vee \exists x \exists y(Rxy \wedge \neg \exists z(Rxz \wedge Rzy) \wedge Px \wedge \neg Py))$$

defines the *finite* linear orderings; a class which is closed under ultrapowers, but not under ultraproducts. Now, this sentence is easily seen to be equivalent to one of the form $\forall P \exists x_1 \dots \exists x_6 \varphi$ with φ as described in 17.8.

In view of lemma 17.1, no improvements upon 17.7 or 17.8 are to be expected. In fact, the following $\Pi_1^1(R)$-sentence ψ is not preserved under ultrapowers, although it is equivalent to both one of the form $\forall P \exists x_1 \exists x_2 \forall x_3 \forall x_4 \varphi$ and one of the form $\forall P \forall x_1 \forall x_2 \exists x_3 \exists x_4 \varphi$; with φ as described in 17.8. Let the L_0-sentence β axiomatize $Th_{L_0}(\langle IN, <\rangle)$. The sentence

$$\psi = \forall P(\beta \wedge ((\exists x(\neg\exists yRyx \wedge Px) \wedge \forall x(Px \rightarrow \forall y((Rxy \wedge \neg\exists z(Rxz \wedge Rzy)) \rightarrow Py))) \rightarrow \forall xPx))$$

defines $\langle IN, <\rangle$ and is, therefore, not preserved under ultrapowers.

The finite linear orderings form a class which is Σ-elementary, but not elementary. Thus, no equivalent of theorem 8.6 holds here. Still, as was remarked in chapter 8, there are two reduction results:

(i) if a $\Pi_1^1(R)$-sentence is Δ-elementary, then it is elementary;
(ii) if a $\Pi_1^1(R)$-sentence is $\Sigma\Delta$-elementary, then it is Σ-elementary.

Both assertions follow from a routine compactness argument.

Another characterization of *USO1* is obtained by means of a straightforward generalization of the proof of theorem 13.6. Let *UV2* be the set of second-order sentences which are true on all frames.

17.9 *Theorem.* A $\Pi_1^1(R)$-sentence is in *USO1* if and only if it is preserved in passing from a generalized frame in which *UV2* holds to the underlying (standard) frame.

Proof. The proof from left to right is exactly like that of 13.6. For the converse, it will be shown that any $\Pi_1^1(R)$-sentence φ with the mentioned preservation property is preserved under ultraproducts.

Consider a set of frames $\{\mathbf{F}_i \mid i \in I\}$ such that $\mathbf{F}_i \models \varphi$ for each $i \in I$. Let U be any ultrafilter on I. It must be shown that $\mathbf{F} = \Pi_U \mathbf{F}_i \models \varphi$. To see this, a generalized frame $\langle \mathbf{F}, W\rangle$ is constructed in which both φ and *UV2* hold. $\langle \mathbf{F}, W\rangle$ is found through the ultraproduct construction first given in chapter 4 for the case of general frames. W consists of all predicates $\Pi_U\{S_i \mid i \in I\}$, where, for some $n \geqslant 1$, $S_i \subseteq W_i^n$ $(i \in I)$; defined by setting

$$\Pi_U \{S_i \mid i \in I\}(f_U^1, \ldots, f_U^n) \text{ iff } \{i \in I \mid S_i(f^1(i), \ldots, f^n(i))\} \in U.$$

For any formula ψ of the present second-order logic, with the free individual variables $x_1, \ldots, x_k$ and the free predicate variables $X_1, \ldots, X_m$, the following equivalence holds:

$$\langle \mathbf{F}, W\rangle \models \psi[f_U^1, ..., f_U^k; \Pi_U\{S_i^1 \mid i \in I\}, ..., \Pi_U\{S_i^m \mid i \in I\}]$$

if and only if

$$\{i \in I \mid \mathbf{F}_i \models \psi[f^1(i), ..., f^k(i); S_i^1, ..., S_i^m]\} \in U.$$

The proof of this is by induction on the complexity of ψ. The cases $\psi = Rx_ix_j$, $\psi = \neg\chi$, $\psi = \chi_1 \rightarrow \chi_2$ and $\psi = \exists y\chi$ are treated exactly as in the proof of the theorem of Łoś. If ψ is $X_jx_{k_1} \ldots x_{k_{n_j}}$, then the argument is as follows.

$\langle \mathbf{F}, W\rangle \models X_jx_{k_1} \ldots x_{k_{n_j}} [f_U^1, ..., f_U^k; \Pi_U\{S_i^1 \mid i \in I\}, ..., \Pi_U\{S_i^m \mid i \in I\}]$ iff (by the truth definition) $\Pi_U\{S_i^j \mid i \in I\}(f_U^{k_1}, ..., f_U^{k_{n_j}})$ iff (by definition) $\{i \in I \mid S_i^j(f^{k_1}(i), ..., f^{k_{n_j}}(i))\} \in U$ iff (by the truth definition) $\{i \in I \mid \mathbf{F}_i \models X_jx_{k_1} \ldots x_{k_{n_j}}[f^1(i), ..., f^k(i); S_i^1, ..., S_i^m]\} \in U$.

Finally, if ψ is $\exists Y\chi$, then the argument is like that for the individual quantifier. If $\langle \mathbf{F}, W\rangle \models \exists Y\chi[f_U^1, ..., f_U^k; \Pi_U\{S_i^1 \mid i \in I\}, ..., \Pi_U\{S_i^m \mid i \in I\}]$, then some predicate $\Pi_U\{S_i \mid i \in I\}$ exists such that $\langle \mathbf{F}, W\rangle \models \chi[f_U^1, ..., f_U^k; \Pi_U\{S_i^1 \mid i \in I\}, ..., \Pi_U\{S_i^m \mid i \in I\}, \Pi_U\{S_i \mid i \in I\}]$. By the induction hypothesis, $\{i \in I \mid \mathbf{F}_i \models \chi[f^1(i), ..., f^k(i); S_i^1, ..., S_i^m, S_i]\} \in U$. Now, this set is included in $\{i \in I \mid \mathbf{F}_i \models \exists Y\chi[f^1(i), ..., f^k(i); S_i^1, ..., S_i^m]\}$; which is, therefore, in U as well.

Conversely, if the latter set X belongs to U, then predicates S_i may be chosen, for each $i \in X$ (outside of X, the choice is arbitrary), such that $\{i \in I \mid \mathbf{F}_i \models \chi[f^1(i), ..., f^k(i); S_i^1, ..., S_i^m, S_i]\}$ contains X and thus belongs to U. Clearly, then, $\Pi_U\{S_i \mid i \in I\}$ is the required predicate such that $\langle \mathbf{F}, W\rangle \models \chi[f_U^1, ..., f_U^k; \Pi_U\{S_i^1 \mid i \in I\}, ..., \Pi_U\{S_i^m \mid i \in I\}, \Pi_U\{S_i \mid i \in I\}]$.

That $\langle \mathbf{F}, W\rangle$ is a generalized frame follows immediately from the above equivalence. Moreover, since both φ and $UV2$ are true in all factors $\mathbf{F}_i$, they will be true in $\langle \mathbf{F}, W\rangle$. QED.

As for the syntactic complexity of $USO1$, a more definite answer can be given than for $\bar{M}1$ (cf. corollary 7.9).

17.10 *Theorem.* $USO1$ is not arithmetically definable.

Proof. Take any $\Pi_1^1(R)$-sentence φ defining $\langle V_\omega, \in\rangle$ (i.e., the hereditarily finite sets) up to isomorphism. E.g., let φ be ZF^2 without the axiom of infinity; but with an additional axiom stating that the class of ordinals is isomorphic to the "real" ω. Let α be any arithmetical sentence; i.e., $\alpha = \alpha(0, S, +, \bullet)$. $0, S, +$ and $\bullet$ may be defined on the basis of φ (be it only locally), so α has an effective translation α^* in L_0 such that $IN \models \alpha$ if and only if $\varphi \models \alpha^*$. We claim that

$$\varphi \models \alpha^* \text{ if and only if } \varphi \vee \alpha^* \in USO1;$$

where $\varphi \vee \alpha^*$ may be considered to be a $\Pi^1_1(R)$-sentence, obtained by widening the scope of the universal second-order quantifiers in φ to include α^*.

First, suppose that $\varphi \models \alpha^*$. Then, clearly, $(\varphi \vee \alpha^*) \leftrightarrow \alpha^*$ is true on all frames; and so $\varphi \vee \alpha^* \in USO1$. Next, suppose that, for some L_0-sentence β, $(\varphi \vee \alpha^*) \leftrightarrow \beta$ is true on all frames. Now, $\langle V_\omega, \in\rangle \models \varphi$, so $\langle V_\omega, \in\rangle \models \beta$. Let $\mathbf{F}$ be any *uncountable* L_0-elementary extension of $\langle V_\omega, \in\rangle$. $\mathbf{F}$ is not isomorphic to $\langle V_\omega, \in\rangle$, and hence $\mathbf{F} \not\models \varphi$. On the other hand, $\mathbf{F} \models \beta$ and, therefore, $\mathbf{F} \models \varphi \vee \alpha^*$. It follows that $\mathbf{F} \models \alpha^*$ and hence $\langle V_\omega, \in\rangle \models \alpha^*$. In other words, $\varphi \models \alpha^*$. QED.

Suitable choices for φ in the above proof will show that *USO1* is not even arithmetically definable when attention is restricted to $\Pi^1_1(R)$-sentences with only one (unary) second-order quantifier.

But, when the parameter R is dropped (i.e., only identity and predicate variables remain), then *USO1* becomes arithmetically definable.

17.11 *Theorem.* The set of sentences in *USO1* which contain no occurrences of the predicate constant R has an arithmetical definition which is Σ^0_2.

Proof. It suffices to show that $\{\varphi \mid \varphi$ is a $\Pi^1_1(R)$-sentence without occurrences of R for which a first-order pure identity sentence ψ exists such that $\varphi \leftrightarrow \psi$ is true on all domains$\}$ has a Σ^0_2-definition. ("Σ^0_2" refers to the *arithmetical hierarchy* in the usual way.) Now, this follows from two observations.

First, the notion "$\psi \to \varphi$ is true on all domains" is Σ^0_1. Secondly, the notion "$\varphi \to \psi$ is true on all domains" is Π^0_1. For, then, the notion "there exists ψ such that both $\psi \to \varphi$ and $\varphi \to \psi$ are true on all domains" will be Σ^0_2. (Schematically: $\exists(\exists A \wedge \forall B) = \exists\exists(A \wedge \forall B) = \exists\exists\forall(A \wedge B)$.)

The first observation is obvious. $\psi \to \varphi$ is true on all domains iff $\psi \models \varphi^-$ (where φ^- is φ without its second-order quantifiers) and, by the completeness theorem for first-order logic, this notion is Σ^0_1. The second observation requires some calculation. Note, for a start, that a ψ^* can be obtained effectively from ψ such that ψ^* is a logical equivalent of ψ which has one of the forms:

$$\text{(i) } \bot, \quad \text{(ii) } \top, \quad \text{(iii) } (P_{n_1} \vee \ldots \vee P_{n_k}) \text{ or } \neg(P_{n_1} \vee \ldots \vee P_{n_k}).$$

(Here P_i is the pure identity sentence expressing the fact that the domain has precisely i elements.) Thus, the statement "$\varphi \rightarrow \psi$ is true on all domains" becomes equivalent to the disjunction "(i) $\psi^* = \bot$ and $\neg\varphi$ is true on all domains, or (ii) $\psi^* = \top$ and ready, or (iii) $\psi^* = (P_{n_1} \vee \ldots \vee P_{n_k})$ and $\varphi \rightarrow (P_{n_1} \vee \ldots \vee P_{n_k})$ is true on all domains, or (iv) $\psi^* = \neg(P_{n_1} \vee \ldots \vee P_{n_k})$ and $(P_{n_1} \vee \ldots \vee P_{n_k}) \rightarrow \neg\varphi$ is true on all domains". It suffices to show that all four disjuncts have a Π^0_1-definition. Let $\varphi = \forall X_1 \ldots \forall X_n \chi$, where $\chi = \chi(X_1, \ldots, X_n, =)$ is a first-order sentence. (i) amounts to "$\exists X_1 \ldots \exists X_n \neg\chi$ is true on all domains". By the Löwenheim-Skolem-Tarski theorem, $\exists X_1 \ldots \exists X_n \neg\chi$ is true on all infinite domains if it is true on any one infinite domain. Moreover, by the compactness theorem, if $\exists X_1 \ldots \exists X_n \neg\chi$ is true on all finite domains, then it is true on at least one infinite domain. It follows that $\exists X_1 \ldots \exists X_n \neg\chi$ is true on all domains iff it is true on all *finite* domains. The latter notion is Π^0_1. Case (ii) is trivial. Case (iii) is like case (i): it amounts to "$\exists X_1 \ldots \exists X_n \neg\chi$ is true on all finite domains whose cardinality is different from $n_1, .., n_k$". Finally, case (iv) amounts to the assertion that $\exists X_1 \ldots \exists X_n \neg\chi$ holds on the finite domains of the cardinalities $n_1, \ldots, n_k$: something which may be checked recursively. QED.

It was remarked already that generalized frames satisfy closure conditions not unlike the one in the definition of $\bar{M}_1^{def}$ (9.14). Indeed the following notion generalizes the idea underlying that definition.

17.12 *Definition.* *GUSO1* is the set of $\Pi^1_1(R)$-sentences which are preserved in passing from any generalized frame $\langle \mathbf{F}, W \rangle$ in which they hold to the underlying frame $\mathbf{F}$.

A result like theorem 9.15 holds here as well.

17.13 *Theorem.* $GUSO1 \subsetneq USO1$

GUSO1 is recursively enumerable.

Proof. It follows from the proof of theorem 17.9 that $GUSO1 \subseteq USO1$. The example of $(Lp \rightarrow LLp) \wedge (LMp \rightarrow MLp)$ works here too: it belongs to $USO1$, but not to $GUSO1$ (cf. chapter 9). Finally, that $GUSO1$ is recursively enumerable follows from the next result. QED.

17.14 *Theorem.* A $\Pi^1_1(R)$-sentence φ is in $GUSO1$ if and only if there exists some L_0-sentence ψ such that $\varphi \leftrightarrow \psi$ is true on all generalized frames.

Proof. If a sentence ψ as described exists, and if $\langle \mathbf{F}, W \rangle$ is a generalized frame on which φ holds, then also $\langle \mathbf{F}, W \rangle \models \psi$. Since ψ is an L_0-sentence, this means that $\mathbf{F} \models \psi$ and hence also $\mathbf{F} \models \varphi$. (Any frame constitutes a generalized frame in the obvious way.)

If $\varphi \in GUSO1$ as defined above, then, by theorem 17.13, $\varphi \in USO1$. I.e., an L_0-sentence ψ exists such that $\varphi \leftrightarrow \psi$ is true on all frames. As it happens, $\varphi \leftrightarrow \psi$ is even true on all *generalized* frames. For, let $\langle \mathbf{F}, W \rangle \models \varphi$. Since $\varphi \in GUSO1$, $\mathbf{F} \models \varphi$. Then also $\mathbf{F} \models \psi$, and hence $\langle \mathbf{F}, W \rangle \models \psi$; because ψ is an L_0-sentence. If, on the other hand, $\langle \mathbf{F}, W \rangle \models \psi$, then, for the same reason, $\mathbf{F} \models \psi$ and so $\mathbf{F} \models \varphi$. Because φ is a $\Pi^1_1(R)$-sentence, it holds a fortiori that $\langle \mathbf{F}, W \rangle \models \varphi$. QED.

Theorem 17.14 connects $GUSO1$ and $USO1$ in a very simple way. (Unfortunately, the connection between the modal classes $\bar{M}_1^{def}$ and $\bar{M}1$ is not as simple. $\bar{M}_1{}^{def}$ is *not* the class of modal formulas which are equivalent to some L_0-sentence on all *general* frames. That class consists of only the closed modal formulas (cf. lemma 13.2).)

The final results of this chapter are concerned with $\Pi^1_1(R)$-definability of classes of frames. A semantic characterization of these classes will be given, derived from Keisler's results about first-order definable classes of frames; much as theorem 16.4 was derived from Birkhoff's results about equational varieties. The more general question which classes of frames are definable by means of any set of sentences of our second-order language turns out to be easier to answer. The reader is, therefore, invited to skip the next part of this chapter until he/she has read the relevant parts of chapter 18; especially theorem 18.13 and its proof.

The result which will be looked for in the sequel has the form "A class of frames is defined by a set of $\Pi^1_1(R)$-sentences if and only if it is closed under ... (*suitable semantic operations*)." Now it is easy to see that, for a given class **K** of frames, the smallest $\Pi_1^1(R)$-definable class of frames to contain **K** is $MOD(Th_{\Pi_1^1(R)}(\mathbf{K}))$. Here $Th_{\Pi_1^1(R)}(\mathbf{K})$ is the set of $\Pi_1^1(R)$-sentences which are true in all frames in **K**; and, for any set Σ, $MOD(\Sigma)$ is the class of structures in which Σ holds. The idea is to find out if $MOD(Th_{\Pi_1^1(R)}(\mathbf{K}))$ has some reasonable semantic characterization. For the case of first-order logic, there exists a related result. Let $Th_{univ}(\mathbf{K})$ be the set of *universal* first-order sentences which are true in all structures in **K**. Clearly, $MOD(Th_{univ}(\mathbf{K}))$ is the smallest universally definable class of frames to contain **K**. Now let $I(\mathbf{K})$ be the class of *isomorphic images* of structures in **K**, $S(\mathbf{K})$ the class of *substructures* of structures in **K**, and $U(\mathbf{K})$ the class of *ultraproducts* of structures in **K**; for any class **K**. Then the following equation holds.

17.15 *Theorem.* $MOD(Th_{univ}(\mathbf{K})) = ISU(\mathbf{K})$.

Proof. $\supseteq$ follows from the well-known preservation properties of U, S and I. $\subseteq$ follows from the fact that any model of $Th_{univ}(\mathbf{K})$ can be isomorphically embedded in an ultraproduct of members of **K** (cf. Chang & Keisler [17]). QED.

It would be pleasant to have a similar connection here. Say, if $\mathbf{F} \models Th_{\Pi_1^1(R)}(\mathbf{K})$, then $\mathbf{F} \in IS^*U^*(\mathbf{K})$; for some suitable notions S^* and U^* of substructure and ultraproduct. A closure condition on **K** is then easily extracted from the statement "$\mathbf{F} \in IS^*U^*(\mathbf{K})$". It so happens that the actual situation is a little more complicated; but this is still the main idea.

Consider any frame **F** such that $\mathbf{F} \models Th_{\Pi_1^1(R)}(\mathbf{K})$. Just following the procedure in first-order logic (referred to in the proof of 17.15) yields a *many-sorted ultraproduct* (cf. chapter 18) $\mathbf{M} = \Pi_U \mathbf{F}_i$ of structures $\mathbf{F}_i \in \mathbf{K}$ ($i \in I$) such that each $\Pi_1^1(R)$-sentence which is true in **M** is also true in **F**. To see this, consider $\Sigma = \{\neg\varphi \mid \varphi$ is a $\Pi_1^1(R)$-sentence such that $\mathbf{F} \models \neg\varphi\}$. Each finite subset of Σ has a model in **K**. For, suppose otherwise. Then, for some $\neg\varphi_1, \ldots, \neg\varphi_k \in \Sigma, \neg(\neg\varphi_1 \wedge \ldots \wedge \neg\varphi_k)$ (i.e., $\varphi_1 \vee \ldots \vee \varphi_k$) would be true in all structures in **K**. Now, by an easy quantifier manipulation, $\varphi_1 \vee \ldots \vee \varphi_k$ is equivalent to a single $\Pi_1^1(R)$-sentence φ; which is, then, in $Th_{\Pi_1^1(R)}(\mathbf{K})$. This is a contradiction, for such sentences φ hold in **F**, by assumption. A construction as in the well-known ultraproduct proof of the Compactness Theorem yields a many-sorted (!) ultraproduct **M** as required.

Actually, if **M** has the required property with respect to **F**, then so does its *full* ultraproduct. But, in the course of our argument, the use of many-sorted ultraproducts will become inevitable in any case.

The next step is to use **F** and **M** as starting points in an elementary chain construction which will yield chain limits $\mathbf{F}_1$ and $\mathbf{M}_1$, respectively; such that $\mathbf{F}_1$ is isomorphic to a suitable kind of substructure of $\mathbf{M}_1$. (Say $\mathbf{F}_1 \in S^*(\{\mathbf{M}_1\})$.) Then, since $\mathbf{F}(\mathbf{M})$ and $\mathbf{F}_1(\mathbf{M}_1)$ are elementarily equivalent, **F** will be in $EIS^*EU_1(\mathbf{K})$; where, for any class **K** of many-sorted structures, $E(\mathbf{K})$ is the class of structures elementarily equivalent to some structure in **K**, and $U_1(\mathbf{K})$ is the class of many-sorted ultraproducts of structures in **K** (cf. chapter 18). Some further manipulations will show that, then, in fact $\mathbf{F} \in ES^*U_1(\mathbf{K})$; from which the required main theorem can be extracted.

Certain technical difficulties in the elementary chain construction make it necessary to change from $\Pi_1^1(R)$-sentences to a class of sentences which is more tractable from the many-sorted perspective. The sorts are S_0 (*individuals*), S_1 (*unary predicates*), S_2, ..., etc., as in chapter 18.

17.16 *Definition.* A *U-sentence* is any many-sorted sentence constructed from atomic formulas and their negations using $\wedge$, $\vee$, $\forall$ and $\exists$ for (variables of) the sort S_0, and $\forall$ for (variables of) other sorts.

17.17 *Lemma* (*AC*). For any U-sentence (considered as a second-order sentence), there exists a $\Pi_1^1(R)$-sentence equivalent to it on the class of standard models.

Proof. Bring the U-sentence into a prenex normal form. Then apply the conversion principles:

$$\begin{array}{lll} & \forall x \forall Y \varphi & \leftrightarrow \quad \forall Y \forall x \varphi, \\ (\mathrm{AC}) & \exists x \forall Y \varphi & \leftrightarrow \quad \forall Y' \exists x \varphi' \text{ (cf. chapter 18)} \end{array}$$

to move all universal quantifiers over sorts different from S_0 to the front. QED.

Again, consider any frame **F** such that $\mathbf{F} \models Th_{\Pi_1^1(R)}(\mathbf{K})$. By lemma 17.17, it may as well be supposed that $\mathbf{F} \models Th_U(\mathbf{K})$; where $Th_U(\mathbf{K})$ is the set of all U-sentences which are true in all structures in **K**. The same argument as before yields a structure **M** in $U_1(\mathbf{K})$ such that each U-sentence which is true in **M** is also true in **F**. Now, the construction of elementary chains can start. In order to state the relevant result, it becomes necessary to specify what was meant by S^* in the preceding (partly informal) discussion.

17.18 *Definition.* A *predicate substructure* (the terminology is *ad hoc*) $\mathbf{M}$ of a many-sorted structure $\mathbf{N}$ is a many-sorted substructure of $\mathbf{N}$ with the same domain of individuals (i.e., S_0) as $\mathbf{N}$. For any class $\mathbf{K}$, $S(\mathbf{K})$ is the class of predicate substructures of structures in $\mathbf{K}$.

17.19 *Lemma.* If each U-sentence which is true in some many-sorted structure $\mathbf{M}$ is also true in the many-sorted structure $\mathbf{N}$, then elementary extensions $\mathbf{M}^+$ and $\mathbf{N}^+$ of $\mathbf{M}$ and $\mathbf{N}$ (respectively) exist such that $\mathbf{N}^+$ is isomorphic to a predicate substructure of $\mathbf{M}^+$.

Proof. The argument is rather like the ones used in chapter 15. The starting position is that of two structures $\mathbf{M}$, $\mathbf{N}$ for some language L, such that $\mathbf{M} - U(L) - \mathbf{N}$; i.e., each U-sentence of L which is true in $\mathbf{M}$ is also true in $\mathbf{N}$. Now, step by step, an "upper" elementary chain $\mathbf{M}$, $\mathbf{M}_1$, $\mathbf{M}_2$, ... and a "lower" chain $\mathbf{N}$, $\mathbf{N}_1$, $\mathbf{N}_2$, ... will be constructed. The first step is typical for the procedure as a whole.

For each w in $\mathbf{M}$'s domain of *individuals*, add a unique new constant w to the language L, to obtain a language L'. Expand $\mathbf{M}$ to an L'-structure $\mathbf{M}'$ by interpreting each w as w. Now each finite subset of $\Sigma = \{\varphi \mid \varphi$ is a U-sentence in L' which is true in $\mathbf{M}'\}$ has a model which is an expansion of $\mathbf{N}$. For, let $\varphi_1, \ldots, \varphi_k \in \Sigma$, containing (in all) the new constants $\mathrm{w}_1, \ldots, \mathrm{w}_m$. Then, for variables $x_1, \ldots, x_m$ of the sort S_0 not occurring in $\varphi_1, \ldots, \varphi_k$, $\mathbf{M} \models \exists x_1 \ldots \exists x_m[x_1/\mathrm{w}_1, \ldots, x_m/\mathrm{w}_m](\varphi_1 \wedge \ldots \wedge \varphi_k)$, and hence this U-sentence of L(!) holds in $\mathbf{N}$. It follows that Σ has a model $\mathbf{N}'_1$ which is an L-elementary extension of $\mathbf{N}$. Note that $\mathbf{M}' - U(L') - \mathbf{N}'_1$.

Now, for each w in *all domains* of $\mathbf{N}'_1$, add a new individual constant w to L'; to obtain a language L_1. Expand $\mathbf{N}'_1$ to an L_1-structure $\mathbf{N}_1$ by interpreting each w as w. Each finite subset of $\Delta = \{\neg\varphi \mid \varphi$ is a U-sentence in L_1 such that $\mathbf{N}_1 \models \neg\varphi\}$ has a model which is an expansion of $\mathbf{M}'$. For, let $\neg\varphi_1, \ldots, \neg\varphi_k \in \Delta$, containing (in all) the new $(L_1 - L')$-constants $\mathrm{w}_1, \ldots, \mathrm{w}_m$. Suppose that no expansion of $\mathbf{M}'$ satisfies $\neg\varphi_1 \wedge \ldots \wedge \neg\varphi_k$. Then the U-sentence $\forall x_1 \ldots \forall x_m[x_1/\mathrm{w}_1, \ldots, x_m/\mathrm{w}_m](\varphi_1 \vee \ldots \vee \varphi_k)$ of L' holds in $\mathbf{M}'$; whence it must hold in $\mathbf{N}'_1$, which is a contradiction. It follows that Δ has a model $\mathbf{M}_1$ which is an L'-elementary extension of $\mathbf{M}'$ such that $\mathbf{M}_1 - U(L_1) - \mathbf{N}_1$. Etcetera.

structures:

$$\begin{tikzcd}[column sep=small]
\mathbf{M}, \arrow[d] & \mathbf{M}' \arrow[dr] & \prec & \mathbf{M}_1, \arrow[d] & \mathbf{M}'_1 \arrow[dr] & & \\
\mathbf{N} & \prec & \mathbf{N}'_1, & \mathbf{N}_1 & \prec & \mathbf{N}'_2, & \mathbf{N}_2
\end{tikzcd}$$

languages: L, L', L', L_1, L'_1, L'_1, L_2, ...

Finally, consider the respective chain limits $\mathbf{N}^+$ and $\mathbf{M}^+$. An obvious isomorphism embeds $\mathbf{N}^+$ in $\mathbf{M}^+$. Note that, by the choice of the sets Σ, the isomorphism is only onto with respect to the domain of *individuals* of $\mathbf{M}^+$. In other words, $\mathbf{N}^+$ is isomorphic to a predicate substructure of $\mathbf{M}^+$. QED.

The method of proof of lemma 17.19 applies more generally to the case with any number of constant domains; instead of just that for the first sort S_0. Definition 17.16 has to be adapted accordingly, of course. Moreover, the proof is easily extended to show that preservation under predicate substructures and definability by means of U-sentences coincide.

Application of lemma 17.19 to the previous $\mathbf{M}$ and $\mathbf{F}$ shows that some elementary equivalent $\mathbf{F}^+$ of $\mathbf{F}$ is isomorphic to some predicate substructure of an elementary equivalent $\mathbf{M}^+$ of some structure $\mathbf{M}$ in $U_1(\mathbf{K})$. Put more briefly, $\mathbf{F} \in EIS^*EU_1(\mathbf{K})$.

One obvious simplification can be made. For any class $\mathbf{K}$, if $\mathbf{F} \in EI(\mathbf{K})$, then $\mathbf{F} \in E(\mathbf{K})$; and hence $\mathbf{F} \in ES^*EU_1(\mathbf{K})$. A further simplification requires some more thought.

17.20 *Lemma.* For any class $\mathbf{K}$ of structures, $ES^*EU_1(\mathbf{K}) \subseteq ES^*U_1(\mathbf{K})$.

Proof. Let $\mathbf{M} \in ES^*EU_1(\mathbf{K})$. I.e., there exists an elementary equivalent $\mathbf{M}_1$ of $\mathbf{M}$ which is a predicate substructure of some $\mathbf{M}_2$ which — in its turn — is elementarily equivalent to some many-sorted ultra-product $\Pi_U\, \mathbf{M}_i$ of structures $\mathbf{M}_i$ ($i \in I$) in $\mathbf{K}$.
In a picture,

$$\mathbf{M} \equiv \mathbf{M}_1 \subseteq \mathbf{M}_2 \equiv \Pi_U\, \mathbf{M}_i.$$

Now, by lemma 18.10, $\mathbf{M}_2$ and $\Pi_U\mathbf{M}_i$ have isomorphic many-sorted ultrapowers, say $\Pi_V\mathbf{M}_2$ and $\Pi_V\Pi_U\mathbf{M}_i$. By lemma 18.11, the latter structure belongs to $IU_1(\mathbf{K})$ and, then, so does $\Pi_V\mathbf{M}_2$. Also the following obvious conversion principle holds:

If $\mathbf{M}_i$ is a predicate substructure of $\mathbf{N}_i$ for each $i \in I$, and if U is an ultrafilter on I, then the many-sorted ultraproduct $\Pi_U\mathbf{M}_i$ is isomorphic to a predicate substructure of the many-sorted ultraproduct $\Pi_U\mathbf{N}_i$. (Use the natural embedding.)

Therefore, the many-sorted ultrapower $\Pi_V\mathbf{M}_1$ is isomorphic to a predicate substructure of $\Pi_V\mathbf{M}_2$ and it belongs, then, to $IS^*IU_1(\mathbf{K})$. Clearly, for any class $\mathbf{K}$, $S^*I(\mathbf{K}) \subseteq IS^*(\mathbf{K})$; and hence $\Pi_V\mathbf{M}_1$ belongs to $IIS^*U_1(\mathbf{K})$, i.e. to $IS^*U_1(\mathbf{K})$. Moreover, $\mathbf{M}_1$ is elementarily equivalent to $\Pi_V\mathbf{M}_1$; which implies that $\mathbf{M} \in EEIS^*U_1(\mathbf{K})$, $\mathbf{M}$ being elementarily equivalent to $\mathbf{M}_1$. From this, the conclusion follows at once: $\mathbf{M} \in ES^*U_1(\mathbf{K})$. QED.

The upshot of the preceding discussion may be codified as follows.

17.21 *Lemma.* If $\mathbf{K}$ is a class of frames and $\mathbf{F}$ is a frame such that $\mathbf{F} \models Th_{\Pi^1_1(R)}(\mathbf{K})$, then $\mathbf{F} \in ES^*U_1(\mathbf{K})$.

In other words, if $\mathbf{F}$ belongs to $MOD(Th_{\Pi^1_1(R)}(\mathbf{K}))$, then $\mathbf{F}$ is elementarily equivalent to some predicate substructure $\mathbf{M}$ of a many-sorted ultraproduct $\Pi_U\mathbf{F}_i$ of structures $\mathbf{F}_i (i \in I)$ in $\mathbf{K}$. Again by lemma 18.10, $\mathbf{F}$ and $\mathbf{M}$ will have isomorphic many-sorted ultrapowers $\Pi_V\mathbf{F}$ and $\Pi_V\mathbf{M}$. By the above conversion principle, $\Pi_V\mathbf{M}$ is isomorphic to some predicate substructure of the ultrapower $\Pi_V\Pi_U\mathbf{F}_i$; which is itself isomorphic to some many-sorted ultraproduct $\Pi_S\mathbf{G}_j$ of frames $\mathbf{G}_j$ in $\mathbf{K}$. Put in terms of the above symbolism, $\Pi_V\mathbf{F} \in IIS^*IU_1(\mathbf{K})$, or — using obvious simplifications — $\Pi_V\mathbf{F} \in IS^*U_1(\mathbf{K})$. This is the final result. It remains to coin some terminology to describe it.

17.22 *Definition.* $\mathbf{M}$ is a *many-sorted subultraproduct* of a set of structures $\{\mathbf{M}_i \mid i \in I\}$ if $\mathbf{M}$ is a predicate substructure of some many-sorted ultraproduct $\Pi_U\mathbf{M}_i$ (where U is an ultrafilter on I).

Recall the concept of "ultraroot"; cf. definition 18.12.

17.23 *Theorem.* A class **K** of standard models is defined by a set of $\Pi^1_1(R)$-sentences if and only if it is closed under frames which are ultraroots of many-sorted subultraproducts of sets of frames $\{\mathbf{F}_i \mid i \in I\}$ in **K**.

This theorem is heavily "proof-generated", of course (cf. the similar remark made about theorem 16.4); but the argument leading up to it may still present some independent interest.

CHAPTER XVIII

SECOND-ORDER LOGIC

The second-order language of this chapter is the same as that of the preceding one. The difference is that *all* sentences will be considered here, not just the $\Pi^1_1(R)$-sentences. Only very few results are known about this kind of logic. Most of these are obtained by stretching first-order methods to the utmost of their capacity: which still does not take one far into this area. Second-order model theory is yet to be developed, as was rightly noted in Chang & Keisler [17]. The present chapter does not pretend to remedy this; having been added only as a kind on natural sequel in the chain of generalization from "modal formula" to "$\Pi^1_1(R)$-sentence" to... .

Probably the most striking feature of second-order logic is its heavy dependence upon set-theoretic assumptions about the class of all structures. E.g., axioms of choice are needed for proving even the simplest kind of quantifier-exchange rule, such as $\exists x \forall Y\ \varphi \leftrightarrow \forall Y' \exists x\ \varphi'$; where x is an individual variable, Y an n-ary predicate variable and Y' a new $n+1$-ary predicate variable; while φ' is obtained from φ by replacing subformulas of the form $Yz_1...z_n$ by $Y'z_1...z_nx$. Such principles are needed to prove (*prenex*) *normal form* theorems to the effect that any second-order sentence is equivalent to one of the form "sequence of second-order quantifiers - sequence of first-order quantifiers - quantifier-free matrix".

Another example are *Skolem* reductions such as $\forall x \exists y Rxy \rightarrow \exists f \forall x Rxfx$.

In contrast, first-order logic is relatively free from such matters. Still, in one of the theses attached to the original dissertation [5], it was remarked, *à propos* of the paper "Relativity Phenomena in Set Theory", «Synthese» 27 (1974), pp.189-198, by C. Åberg, that there are extensions of Zermelo-Fraenkel set theory in which more L_0-sentences are provably universally valid than there are L_0-logical truths; but this is a mere curiosity.

For readers interested in metaphysics, here is the argument. The following equivalence cannot hold for all L_0-sentences α:

(★) $\not\models \alpha$ if and only if $ZF \vdash$ "$\exists \mathbf{F}\ \mathbf{F} \not\models \alpha$".

For, otherwise, the set of non-universally valid L_0-sentences would be recursively enumerable. But, then, the set of universally valid L_0-sentences, which is recursively enumerable by the completeness theorem, would be recursive — by Post's theorem. This contradicts the theorem of Church, however. Now, it is easy to see that one half of (★) is correct: if $ZF \vdash$ "$\exists \mathbf{F}\ \mathbf{F} \not\models \alpha$", then $\not\models \alpha$. For, if $\models \alpha$, then $\vdash \alpha$ and hence, clearly, $ZF \vdash$ "$\forall \mathbf{F}\ \mathbf{F} \models \alpha$"; by a direct syntactic argument. It follows that $ZF \not\vdash$ "$\exists \mathbf{F}\ \mathbf{F} \not\models \alpha$"; always assuming ZF to be consistent.

The conclusion must be that the other direction of (★) fails. I.e., an L_0-sentence α exists such that $\not\models \alpha$, but also $ZF \not\vdash$ "$\exists \mathbf{F}\ \mathbf{F} \not\models \alpha$"; or, equivalently, such that $ZF \cup \{$"$\forall \mathbf{F}\ \mathbf{F} \models \alpha$"$\}$ is consistent. The latter theory would be an extension of ZF with a "logical truth" α for which, in reality, $\not\models \alpha$.

Note that the following equivalence does hold for all L_0-sentences α:

$\models \alpha$ if and only if $ZF \vdash$ "$\forall \mathbf{F}\ \mathbf{F} \models \alpha$".

Moreover, the usual extensions of ZF do not increase the amount of L_0-logical truths. E.g., if ZFL (= ZF plus the Axiom of Constructibility) $\vdash$ "$\forall \mathbf{F}\ \mathbf{F} \models \alpha$", then, by the completeness theorem for L_0 (which is available inside ZF already), $ZFL \vdash$ "α is L_0-provable". But, the latter assertion is arithmetical; whence — ZFL being conservative over ZF with respect to arithmetical statements — $ZF \vdash$ "α is L_0-provable". Again by the completeness theorem in ZF, then, $ZF \vdash$ "$\forall \mathbf{F}\ \mathbf{F} \models \alpha$".

After this excursion we return to more prosaic subjects.

Standard models are still the original frames. But, we shall now consider more complex general models, whose sets of predicates are closed under the more comprehensive requirement formulated in the following definition.

18.1 *Definition.* A *general model* $\langle \mathbf{F}, W \rangle$ consists of a frame $\mathbf{F}$ together with a set W of predicates on $\mathbf{F}$ such that, for all second-order formulas $\varphi = \varphi(X_1, \ldots, X_k, y_1, \ldots, y_m, z_1, \ldots, z_n)$, all predicates $S_1, \ldots, S_k$ in W (such that S_i is j-ary if X_i is $(1 \leqslant i \leqslant k)$) and all elements $w_1, \ldots, w_m$ of W, the following predicate belongs to W:

$$\{\langle v_1, \ldots, v_n\rangle \mid v_1, \ldots, v_n \in W \text{ and } \langle \mathbf{F}, W\rangle \models \varphi[S_1, \ldots, S_k; w_1, \ldots, w_m; v_1, \ldots, v_n]\}.$$

This closure condition is highly impredicative, of course, in that the formulas φ may contain quantifiers ranging over W. Any frame may be identified with the general model $\langle \mathbf{F}, W\rangle$, where W consists of *all* predicates on $\mathbf{F}$. As in chapter 17, a *smallest* general model $\langle \mathbf{F}, W\rangle$ always exists for a given frame $\mathbf{F}$ (take intersections); but — unlike in chapter 17 — no simple constructive definition of such a smallest model can be produced. The use of general models lies in their providing a complete semantics for a very natural system of second-order deduction; cf. Henkin [33].

On the other hand, universal validity *on frames* becomes a highly interesting notion in second-order logic. As was stated in corollary 2.31, the set of universally valid modal formulas is recursive. It was remarked in chapter 17 that the set of universally valid $\Pi^1_1(R)$-sentences is recursively enumerable (though it is not recursive). If no first-order parameters are allowed except identity, then the set of universally valid $\Sigma^1_1(R)$-sentences is Π^0_1-definable; as was shown during the proof of theorem 17.11. But, universal validity at the next level of the syntactic hierarchy becomes hyper-arithmetical. To see this, recall the result of H. C. Doets mentioned in chapter 1: even universal validity in the theory of finite types (cf. chapter 19) is reducible to universal validity of second-order sentences of the form $\forall R \exists P \varphi(R, P)$; where R is a binary predicate variable, P a unary one and φ is a first-order sentence in R and P. Call second-order sentences of the form "universal second-order quantifiers - existential second-order quantifiers -first-order sentence" Π^1_2-sentences. If the second-order quantifier prefix is reversed, then these are called Σ^1_2-sentences; provided that no first-order predicate constants occur. (In matters of validity, these give rise to additional universal second-order quantifiers in front.)

Two direct examples show how complicated such sentences are.

18.2 *Lemma.* The set of Π^1_2-sentences which are true in all standard models is not arithmetically definable. The set of Σ^1_2-sentences which are true in all standard models is not arithmetically definable.

Proof. A suitable translation of arithmetical sentences into such second-order sentences will prove both assertions; by means of Tarski's Undefinability Theorem. The relevant arguments will be sketched now.

First, let PA^2 be the universal second-order sentence (with first-order parameters 0, S, $+$, $\cdot$) axiomatizing second-order *Peano Arithmetic*. PA^2 defines the standard model $\langle IN, 0, S, +, \cdot\rangle$ up to isomorphism. It may be written as $\forall X\alpha \wedge \beta$, where $\forall X\alpha$ is the second-order axiom of induction and β the conjunction of the remaining (first-order) axioms.

For any arithmetical (first-order) sentence γ, it holds that

(i) $\langle IN, 0, S, +, \cdot\rangle \models \gamma$ iff $\forall 0\forall S\forall +\forall\cdot(PA^2 \rightarrow \gamma)$ is true in all standard models.

Now, the latter sentence is equivalent to $\forall 0\forall S\forall +\forall\cdot\exists X((\alpha \wedge \beta) \rightarrow \gamma)$. The first-order quantifier $\forall 0$ may be interchanged with $\forall S$,$\forall +$ and $\forall\cdot$ and also with $\exists X$; because of the reduction principle

$$\forall x\exists Y\varphi \leftrightarrow \exists Y'\forall x\varphi',$$

which is the dual of the one mentioned at the beginning of this chapter. The result is a Π^1_2-sentence on the right-hand side of the equivalence. (Of course, the functional quantifiers have to be replaced by suitable predicate quantifiers: but, these are trivia.)

The equivalence needed to prove the second assertion of the lemma is

(ii) $\langle IN, 0, S, +, \cdot\rangle \models \gamma$ iff $\exists f$(f is a 1-1 function from the domain into the domain which is not onto) $\rightarrow \exists N\exists 0\exists S\exists +\exists\cdot\exists x((PA^2)^N \wedge (\gamma)^N)$ is true in all standard models.

(Here N is a unary predicate variable, and $(PA^2)^N$, $(\gamma)^N$ are the obvious relativizations.)

The consequent sentence of the formula on the right-hand side is equivalent to a Σ^1_2-sentence; by considerations similar to the above. Therefore, the whole sentence is equivalent to a Σ^1_2-sentence; by means of quantifier changes according to the schema $\exists \rightarrow \exists\forall$, $\exists(\exists \rightarrow \forall)$, $\exists\forall(\exists \rightarrow -)$, $\exists\forall\forall(- \rightarrow -)$. QED.

No system of deduction will be given here. For the following remarks, it suffices to fix any reasonable notion $\vdash$ of derivability for the present language, complete with respect to the above class of general models. It turns out that the concept of "completeness" — in the sense of definition 6.4 — has no interesting generalization here. To be sure, one can define an analogue.

18.3 *Definition.* A second-order sentence φ is *complete* if, for all second-order sentences ψ, $\varphi \models \psi$ (i.e., $\varphi \rightarrow \psi$ is true in all standard models) if and only if $\varphi \vdash \psi$.

But, this notion turns out to be too strong. Let *UV* be the set of second-order sentences which are true in all standard models.

18.4 *Lemma.* A second-order sentence φ is complete if and only if $\varphi \vdash UV$.

Proof. If φ is complete, then, for any $\psi \in UV$, $\varphi \vdash \psi$; because $\varphi \models \psi$. If, on the other hand, $\varphi \vdash UV$, and if, moreover, $\varphi \models \psi$ (i.e.,$\varphi \rightarrow \psi \in UV$), then $\varphi \vdash \varphi \rightarrow \psi$ and hence $\varphi \vdash \psi$. Moreover, $\vdash$ is chosen so as to ensure that if $\varphi \vdash \psi$, then $\varphi \models \psi$. It follows, then, that φ is complete. **QED.**

Again, there are two notion of first-order definability, one in terms of standard models, the other in terms of general models.

18.5 *Definition.* *SO1* is the set of second-order sentences φ for which an L_0-sentence α exists such that $\varphi \leftrightarrow \alpha$ is true in all standard models. (Cf. definition 17.6.)

GSO1 is the set of second-order sentences φ for which an L_0-sentence α exists such that $\varphi \leftrightarrow \alpha$ is true in all general models. (Cf. theorem 17.14.)

Clearly, *GSO1* $\subseteq$ *SO1*, but not conversely. Moreover, *GSO1* is recursively enumerable; because truth in all general models is a recursively axiomatizable notion (cf. earlier remarks, or Henkin [33]).

Several methods may be applied to obtain a semantic characterization of *SO1*. The first uses Keisler's theorem about elementary classes.

18.6 *Theorem.* A second-order sentence φ is in *SO1* if and only if both φ and $\neg\varphi$ are preserved under ultraproducts.

A second method copies the proof of theorem 17.9.

18.7 *Theorem.* A second-order sentence φ is in *SO1* if and only if both φ and $\neg\varphi$ are preserved in passing from a general model in which *UV* holds to the underlying standard model.

A third method uses Lindström's Theorem characterizing first-order logic in terms of its compactness and Löwenheim-Skolem properties. Let $L1(\varphi)$ be the first-order logic consisting of ordinary predicate logic with at least infinitely many binary predicate constants together with φ as an additional propositional constant, closed under Boolean operators, while admitting relativization to unary predicates.

Some care is needed: addition of φ to just L_0 will not do. For one thing, Lindström's proof does not apply to first-order logics with only a *finite* vocabulary; and, for another, his theorem even fails in such a case. (This was another one of the theses added to the original dissertation [5].) E.g., let φ define the countably infinite frame satisfying $\forall x \forall y Rxy$. φ is not definable in L_0; i.e., the language obtained from L_0 by adding φ and then closing under the Boolean operators is a proper extension of it. And yet it satisfies both the compactness and the Löwenheim-Skolem properties; as is easily checked. (More information on this, and related issues from this chapter may be found in the chapter on "Higher-Order Logic", written by H. C. Doets and the present author, in volume I of the *Handbook of Philosophical Logic*, D. Gabbay & F. Guenthner (eds), Reidel, Dordrecht, 1983.)

Nevertheless, the following result may be formulated.

18.8 *Theorem.* A second-order sentence φ is in *SO1* if and only if $L1(\varphi)$ satisfies the compactness and the Löwenheim-Skolem properties (under the normal semantic interpretation for φ).

There are still other methods available; e.g., one using Fraïssé's Theorem.

A semantic characterization for *GSO1* may be derived from that for *SO1* given in theorem 18.7. The proof is just like that of theorem 17.14.

18.9 *Theorem.* A second-order sentence φ is in *GSO1* if and only if both φ and $\neg\varphi$ are preserved in passing from a general model to the underlying standard model.

This theorem answers a question of S. Orey in [62].

Finally, matters of definability will be treated, like the ones in chapter 16. The main question is when a class of standard models is definable by means of a set of second-order sentences. Some sort of answer will be obtained through the use of Keisler's characterization of elementary classes.

First, consider the case of ordinary first-order logic. Let **K** be any class of structures. Define $Th(\mathbf{K})$ as $\{\varphi \mid \varphi$ is a first-order sentence which is true in all structures in $\mathbf{K}\}$. It is easily seen that the smallest Δ-elementary class of structures to contain **K** is $MOD(Th(\mathbf{K}))$; where $\mathbf{MOD}(\Sigma) =_{\mathrm{def}}$ the class of all structures in which Σ holds. Now, a well-known model-theoretic argument characterizes this class semantically as $IE_sU(\mathbf{K})$. Here, for any class **K** of structures,

$I(\mathbf{K})$ is the class of *isomorphic images* of structures in **K**,
$E_s(\mathbf{K})$ is the class of *elementary substructures* of structures in **K**, and
$U(\mathbf{K})$ is the class of *ultraproducts* of subsets of **K**.

(The point is that any structure in which $Th(\mathbf{K})$ holds is isomorphic to an elementary substructure of some ultraproduct of structures of **K**.) This result goes through for any *many-sorted* language (cf. Enderton [20]), in which the language has different sorts of individual variables ranging over disjoint sorts of domains in the structures.

Now, our second-order language may be regarded as a many-sorted first-order language in an obvious way. (Cf. Enderton [20] or J. D. Monk, *Mathematical Logic*, Springer, Berlin, 1975.) There will be sorts S_0 (for *individuals*), S_1 (for *unary predicates*), S_2 (for *binary predicates*), ..., etc. In structures $\langle \mathbf{F}, W\rangle$, S_0 denotes W, S_1 the unary predicates in W, etc. But, many other structures for this many-sorted language exist with arbitrary domains for the different sorts. There are ways to "normalize" such structures; by requiring certain principles of extensionality to hold. E.g., let E be a binary predicate between objects of S_0 and objects of S_1. In structures where $\forall x_1 \forall x_1'(\forall x_0(Ex_0x_1 \leftrightarrow Ex_0x_1') \rightarrow x_1 = x_1')$ holds, objects s in S_1 may be identified with $\{w \in S_0 \mid E(w, s)\}$. Moreover, one could require a many-sorted version of the Comprehension Principle to hold, in order to secure a better correspondence with general models as defined in definition 18.1. Such requirements are not essential to the following arguments, however.

Many-sorted ultraproducts can be defined in a straightforward manner. When general models are regarded as many-sorted structures in the way indicated above, the ultraproducts defined in the proof of theorem

17.9 become many-sorted ultraproducts of this kind; which have been "normalized" back to general models.

For a class **K** of many-sorted structures, let $U_1(\mathbf{K})$ be the class of many-sorted ultraproducts of structures in **K**. To obtain the main result below, it suffices to consider this notion; together with $E(\mathbf{K})$, defined as the class of structures which are elementarily equivalent (in the many-sorted sense) to some structure in **K**. Instead of the above equation, the following simpler one is needed:

The smallest Δ-elementary class to contain a given class **K** of many-sorted structures is $EU_1(\mathbf{K})$.

The standard proof of this statement uses the fact that any structure in which the many-sorted theory of **K** holds is elementarily equivalent to some (many-sorted) ultraproduct of members of **K**.

Now, the result we are looking for is of the form: "A class of frames is defined by a set of second-order sentences if and only if it is closed under ... (*suitable semantic operations*)". These operations are to be formulated in terms of *frames*, but we will allow ourselves an excursion into a domain of auxiliary structures: just as we did in the proof of theorem 16.4.

If the class **K** is defined by any set of second-order sentences at all, then it is defined by its own second-order theory $Th_2(\mathbf{K})$. Clearly, $\mathbf{K} \subseteq FR(Th_2(\mathbf{K}))$; so the main problem is to make sure that $FR(Th_2(\mathbf{K})) \subseteq \mathbf{K}$.

Now, let $\mathbf{F} \models Th_2(\mathbf{K})$. A connection is to be found between **F** and structures in **K** which can be translated into a closure condition as required above. To find one, the structures will be regarded as many-sorted structures, and $Th_2(\mathbf{K})$ as the many-sorted theory of **K**. Then the above equation becomes applicable: $\mathbf{F} \in EU_1(\mathbf{K})$. (We do not provide new notation to emphasize the change in point of view, for fear of overloading the exposition.)

The notion E is rather artificial, being heavily dependent upon the *language* being used. Fortunately, there exists a way to get rid of it, by using Keisler's well-known theorem stating that any two elementarily equivalent structures have isomorphic ultrapowers. Of course, it must be checked if this theorem also applies to many-sorted logic.

18.10 *Lemma.* If two many-sorted structures have the same many-sorted theory, then they have isomorphic many-sorted ultrapowers.

Proof. One way to see this is by generalizing Keisler's proof. A rather easier method is by finding a way to apply the original (one-sorted) result directly. Now, if the language has only *finitely* many sorts $S_1, \ldots, S_k$, then it is easy. Let $\mathbf{M}_1$, $\mathbf{M}_2$ be any two elementarily equivalent many-sorted structures. Transform the language into a *one-sorted* one by adding unary predicate constants $S_1, \ldots, S_k$. Many-sorted formulas may then be transcribed into formulas of the new language; essentially by replacing quantifiers $\forall x_i$ ranging over the i-th sort by $\forall x(S_i x \rightarrow$ $(1 \leqslant i \leqslant k)$. The exact procedure is in the common text books. Similarly, $\mathbf{M}_1$, $\mathbf{M}_2$ may be expanded to structures $\mathbf{M}_1^1$, $\mathbf{M}_2^1$ for this new one-sorted language. If $\mathbf{M}_1^1$ and $\mathbf{M}_2^1$ can be shown to be elementarily equivalent in the ordinary sense, then Keisler's original theorem applies, and $\mathbf{M}_1^1$ and $\mathbf{M}_2^1$ have isomorphic ultrapowers; from which isomorphic many-sorted ultrapowers for $\mathbf{M}_1$ and $\mathbf{M}_2$ are easily extracted.

To see that $\mathbf{M}_1^1$ and $\mathbf{M}_2^1$ are elementarily equivalent, consider the theory T consisting of the one-sorted transcription of the many-sorted theories of $\mathbf{M}_1$ (and $\mathbf{M}_2$) together with the principles

(i) $\forall x(S_1 x \vee \ldots \vee S_k x)$
(ii) $\forall x(S_i x \rightarrow \neg S_j x) \qquad (1 \leqslant i \neq j \leqslant k)$
(iii) $\forall x_1 \ldots \forall x_m(P x_1 \ldots x_m \rightarrow S_{n_1} x_1 \wedge \ldots \wedge S_{n_m} x_m))$;
for predicates P whose "sortal type" is $\langle n_1, \ldots, n_m \rangle$.

For any one-sorted sentence φ, there exists a one-sorted sentence φ^*, which is the transcription of a many-sorted sentence, such that $\mathrm{T} \vdash \varphi \leftrightarrow \varphi^*$. From this, the above assertion follows at once; by the fact that $\mathbf{M}_1$ and $\mathbf{M}_2$ are elementarily equivalent in the many-sorted sense.

φ^* is obtained from φ by simply relativizing quantifiers. E.g., $\forall x \psi(x)$ becomes $\forall x(S_1 x \rightarrow \psi) \wedge \ldots \wedge \forall x(S_k x \rightarrow \psi)$.

If the language has *infinitely* many sorts $S_1, S_2, \ldots$, then the construction of φ^* becomes more complicated. Here is the principle.

(i) Rewrite φ to an equivalent sentence involving only $=, \neg, \rightarrow$ and $\forall$ as logical primitives. Change bound variables so as to ensure that no two quantifiers have the same bound variable. Let $S_1, \ldots, S_k$ be the sortal constants which either occur in φ or correspond to a sort occurring in some sortal type of a predicate constant occurring in φ.

(ii) Replace, starting from the outside, subformulas of the form $\forall x \alpha$ by $\prod_{i=1}^{k} \forall x(S_i x \rightarrow \alpha) \wedge \forall x(\prod_{i=1}^{k} \neg S_i x \rightarrow \alpha)$. Abbreviate "$\prod_{i=1}^{k} \neg S_i x$" by "$S_c x$".

Again, all bound variables are to be different; which makes it possible to define *the sort of* a bound variable according to the sortal relativization of *its* quantifier.

(iii) Consider all atomic subformulas α one by one. If $\alpha = S_i x$ and the sort of x is different from S_i, then replace α by the *falsum* $\perp$. If $\alpha = Px_1 \ldots x_m$ and any x_i $(1 \leqslant i \leqslant m)$ has a sort different from the one it should have according to the sortal type $\langle n_1, \ldots, n_m \rangle$ of P, then replace α again by $\perp$. Finally, if α is $x_1 = x_2$ and x_1, x_2 have different sorts, then replace α by $\perp$. These replacements yield a formula equivalent (in T) to the original one; in which quantifiers of the form $\forall x(S_c x \rightarrow$ bind only variables occurring in atomic subformulas of the form $x = y$ or $y = x$; where both x and y have the same sort S_c.

(iv) Formula induction shows that any formula of the kind obtained in step (iii) is equivalent to a Boolean combination of formulas in which S_c does not occur and formulas whose only quantifiers are of the form $\forall x(S_c x \rightarrow$ and whose only atomic subformulas are of the forms $x = y$ or $\perp$. It follows that the original sentence φ is equivalent (in T) to a Boolean combination of sentences which are transcriptions of many-sorted sentences and identity sentences about S_c. The latter sentences have definite truth values on the basis of T, and may, therefore, be dropped in an obvious fashion. This yields the required sentence φ^*. QED.

Recall that $\mathbf{F}$ was in $EU_1(\mathbf{K})$. Say that it is elementarily equivalent to a many-sorted ultraproduct $\Pi_U \mathbf{F}_i$ of members $\mathbf{F}_i (i \in I)$ of $\mathbf{K}$. By lemma 18.10, there exists a set J and an ultrafilter V on J such that $\Pi_V \mathbf{F}$ is isomorphic to $\Pi_V \Pi_U \mathbf{F}_i$. Now the latter structure may be simplified a little.

18.11 *Lemma.* For any class $\mathbf{K}$ of many-sorted structures,

$$U_1 U_1(\mathbf{K}) \subseteq IU_1(\mathbf{K}).$$

Proof. Let $\mathbf{F}_j = \Pi_{U_j} \mathbf{F}_{ij}$ $(i \in I_j;\ j \in J)$ be many-sorted ultraproducts of structures in $\mathbf{K}$. Let V be an ultrafilter on J. Make the sets I_j disjoint in some fashion, obtaining ultraproducts $\mathbf{F}'_j$ isomorphic to the original ultraproducts $\mathbf{F}_j$. Next, define a new ultrafilter V^* on $\bigcup_{j \in J} I_j$ by setting $X \in V^*$ iff $\{j \in J \mid X \cap I_j \in U_j\} \in V$. The ultraproduct $\Pi_{V^*} \mathbf{F}'_{ij}$ is isomorphic to $\Pi_V \Pi_{U_j} \mathbf{F}_{ij}$ via the obvious mapping. QED.

Actually, this result holds also for *ordinary* ultraproducts — by the same argument.

By lemma 18.11, the above ultraproduct $\Pi_V\mathbf{F}$ is isomorphic to some many-sorted ultraproduct $\Pi_U\mathbf{F}_i$ of structures $\mathbf{F}_i$ in $\mathbf{K}$. It remains to formulate this result more concisely.

18.12 *Definition.* A many-sorted structure $\mathbf{M}_1$ is an *ultraroot* of a structure $\mathbf{M}_2$ if $\mathbf{M}_2$ is isomorphic to a many-sorted ultrapower of $\mathbf{M}_1$. A many-sorted structure $\mathbf{M}_1$ is an *ultrameans* of a set of such structures $\{\mathbf{M}_i \mid i \in I\}$ if it is an ultraroot of some many-sorted ultraproduct of that set.

The arithmetical analogy will be clear. The above chain of reasoning can now be summarized as

18.13 *Theorem.* A class of standard models is defined by a set of second-order sentences if and only if it is closed under frames which are ultrameans of sets of its members.

It might be possible to obtain analogues of theorem 18.13 by algebraic methods, using the *cylindric algebras* (and some appropriate Birkhoff-type result) of the Tarski School; but — due to the technical difficulties involved — this matter has not been explored here.

CHAPTER XIX

THE THEORY OF FINITE TYPES

This last chapter forms a kind of logical conclusion to the sequence of generalizations starting in chapter 17. It is no more than an appendix, giving the barest outline of Russell's elegant generalization of second-order, or indeed higher-order logic. From many points of view, the theory of finite types L_ω is a very natural system of logic. The technical difficulties it presents are rather formidable, however.

19.1 *Definition.* The set of *types* is the smallest set to contain the *basic types* e (for "entity") and t (for "truth value") and be closed under the formation of ordered couples.

The language L_ω has infinitely many variables and constants of each type. Its terms are defined by recursion.

19.2 *Definition.* The set $TERM_a$ of *terms of type a* is defined according to the clauses, for all types a and b,

(i) the variables and constants of type a belong to $TERM_a$,
(ii) if $\alpha \in TERM_{\langle a,b \rangle}$ and $\beta \in TERM_a$, then $\alpha(\beta) \in TERM_b$ (*functional application*)
(iii) if $\alpha \in TERM_b$ and x is a variable of type a, then $\lambda x.\alpha \in TERM_{\langle a,b \rangle}$ (λ-*abstraction*)
(iv) if α, $\beta \in TERM_a$, then $(\alpha = \beta) \in TERM_t$ (*identity*).

Thus, L_ω has only three logical primitives, whose meaning will be fixed in the following semantics.

19.3 *Definition.* A *standard model* **M** on a non-empty domain D_e (of individuals) consists of D_e together with an *interpretation* I taking constants of type a into D_a; where D_a is defined by the recursion

(i) D_e is given
(ii) D_t is the set of truth values $\{0,1\}$
(iii) $D_{\langle a,b\rangle}$ is $D_b^{D_a}$; i.e., the set of functions with domain D_a taking their values in D_b.

Note that D_e and I determine the standard model completely.

The standard definition of *truth* (or rather *designation*) is as follows.

19.4 *Definition.* Let f be a function assigning values to (at least) every variable occurring free in the term α such that, to variables x of type a, values $f(x)$ are assigned in D_a. The *value of* α *in* **M** *under* $f(V_{\mathbf{M}}(\alpha, f))$ is defined by the recursion

(i) $V_{\mathbf{M}}(\alpha, f) = f(\alpha)$ for variables α
(ii) $V_{\mathbf{M}}(\alpha, f) = I(\alpha)$ for constants α
(iii) $V_{\mathbf{M}}(\alpha(\beta), f) = V_{\mathbf{M}}(\alpha, f)(V_{\mathbf{M}}(\beta, f))$
(iv) $V_{\mathbf{M}}(\lambda x.\alpha, f)$, for a variable x of type a and a term α of type b, is the function f from D_a into D_b assigning to each $w \in D_a$ the value $V_{\mathbf{M}}(\alpha, f^x_w)$; where f^x_w is the function which is like f except for the possible difference that w is assigned to x
(v) $V_{\mathbf{M}}((\alpha = \beta), f) = 1$, if $V_{\mathbf{M}}(\alpha, f) = V_{\mathbf{M}}(\beta, f)$; and 0, otherwise.

A clever series of definitions by L. Henkin shows that the other usual logical constants are definable in terms of the above three.

Again, there are general models.

19.5 *Definition.* A *general model* **M** consists of a family of non-empty sets D_a for each type a; where D_e and D_t are given as above and $D_{\langle a,b\rangle} \subseteq D_b^{D_a}$; together with an interpretation I as before. Moreover, the following condition holds. For each term α of type a, $V_{\mathbf{M}}(\alpha, f)$ — when computed as in definition 19.4, for a suitable function f — belongs to D_a.

Note that definition 19.5 is highly impredicative. Cf. also definition 18.1.

There are two ways to look at *ultraproducts* of standard models. One consists in considering domains D_e only, taking the usual first-order ultraproduct, and then viewing the corresponding standard model. The other way generalizes the many-sorted ultraproduct construction of previous chapters in the following fashion.

Let general models $\mathbf{M}^i$ be given for each $i \in I$; as well as some ultrafilter U on I. The *general ultraproduct* $\mathbf{M} = \Pi_U \mathbf{M}^i$ is defined as follows.

All sets D_a will be of the form $\{\bar{f} \mid f$ is a function with domain I whose i-value is in D^i_a $(i \in I)\}$. The mapping $^-$ to be defined presently will satisfy the following *claim*:

"for all types a, if f and g are functions from I into $\bigcup_{i \in I} D^i_a$ whose i-values are in $D^i_a (i \in I)$, then $\bar{f} = \bar{g}$ if and only if $\{i \in I \mid f(i) = g(i)\} \in U$."

(i) D_e: $\bar{f} = \{g \mid g$ is a function with domain I whose i-values are in D^i_e, and $\{i \in I \mid f(i) = g(i)\} \in U\}$.

This is the usual ultraproduct construction, and the claim is obviously satisfied.

(ii) D_t: $\bar{f} = \begin{cases} 1, \text{ if } \{i \in I \mid f(i) = 1\} \in U \\ 0, \text{ otherwise.} \end{cases}$

It is easily seen that the claim holds for this case too.

(iii) $D_{(a,b)}$: Suppose that D_a and D_b have been defined already such that the claim holds. Let f be any function with domain I whose i-values are in $D^i{}_{(a,b)}$. A function $\bar{f}$ from D_a into D_b is defined by setting, for $\bar{g} \in D_a$, $\bar{f}(\bar{g}) = \overline{\lambda i.f(i)(g(i))}$ (where $f(i)(g(i)) \in D_b$).

$^-$ is *well-defined*. For, suppose that $\bar{g}_1 = \bar{g}$. Then, by the inductive hypothesis for D_a, $\{i \in I \mid g_1(i) = g(i)\} \in U$; whence, by the inductive hypothesis for D_b, $\overline{\lambda i.f(i)(g_1(i))} = \overline{\lambda i.f(i)(g(i))}$; i.e., $\bar{f}(\bar{g}) = \bar{f}(\bar{g}_1)$.

Now for the claim in this new case. It is to be shown that $\bar{f} = \bar{f}_1$ iff $\{i \in I \mid f(i) = f_1(i)\} \in U$. First, suppose that $\bar{f} = \bar{f}_1$. This means that, for all $\bar{g} \in D_a$, $\bar{f}(\bar{g}) = \bar{f}_1(\bar{g})$, or $\overline{\lambda i.f(i)(g(i))} = \overline{\lambda i.f_1(i)(g(i))}$; i.e., by the inductive hypothesis for D_b, that $\{i \in I \mid f(i)(g(i)) = f_1(i)(g(i))\} \in U$. Now, suppose that $\{i \in I \mid f(i) = f_1(i)\} \notin U$. Then $\{i \in I \mid f(i) \neq f_1(i)\} \in U$ and x_i may be chosen in D^i_a (wherever possible) such that $f(i)(x_i) \neq f_1(i)(x_i)$; while x_i may be taken to be arbitrary at all other indices. Because of the assumption that $\bar{f} = \bar{f}_1$, it must hold that $\bar{f}(\overline{\lambda i.x_i}) = \bar{f}_1(\overline{\lambda i.x_i})$ and, by the above, $\{i \in I \mid f(i)(x_i) = f_1(i)(x_i)\} \in U$; which is a contradiction with the definition of $\langle x_i \rangle_{i \in I}$. It remains to prove the converse direction.

Suppose that $\{i \in I \mid f(i) = f_1(i)\} \in U$. Then, for any $\bar{g} \in D_a$, $\bar{f}(\bar{g}) = \overline{\lambda i.f(i)(g(i))} = \overline{\lambda i.f_1(i)(g(i))}$ (by the inductive hypothesis for D_b) $= \bar{f}_1(\bar{g})$.

Finally, put $\mathbf{I}(c) = \overline{\lambda i.I^i(c)}$ for constants c. This concludes the construction of $\mathbf{M}$.

The theorem of Łoś may now be proven in the following form.

19.6 *Theorem.* For any first-order ultraproduct $\mathbf{M}$ of a set $\{\mathbf{M}^i \mid i \in I\}$ as described above, any term α and any assignment f to (at least) the free variables in α,

$$V_{\mathbf{M}}(\alpha, f) = \overline{\lambda i.V_{\mathbf{M}}(\alpha, f_i)};$$

where, if f assigns $\bar{g}$ to x, then f_i assigns $g(i)$ to x.

Theorem 19.6 is proven by a standard induction on the construction of α, taking a few precautions as regards substitutability of variables. An important consequence is the *Compactness Theorem.*

19.7 *Theorem.* For any set T of terms of type a such that, for all finite $T_0 \subseteq T$, there exist $\mathbf{M}, f$ with $V_{\mathbf{M}}(\alpha, f) = V_{\mathbf{M}}(\alpha', f)$ for all $\alpha, \alpha' \in T_0$, there also exist $\mathbf{M}, f$ such that $V_{\mathbf{M}}(\alpha, f) = V_{\mathbf{M}}(\alpha', f)$ for all $\alpha, \alpha' \in T$.

Proof. Choose $\mathbf{M}, f$ as described for every finite $T_0 \subseteq T$; say $\mathbf{M}_{T_0}, f_{T_0}$. Take any ultrafilter U on the set of finite subsets of T which contains all sets of the form $\{X \subseteq T \mid X$ is finite and $X \supseteq T_0\}$, and consider $\mathbf{M} = \Pi_U \mathbf{M}_{T_0}$ with the obvious f. If $\alpha, \alpha' \in T$, then $V_{\mathbf{M}}(\alpha, f) = \overline{\lambda T_0.V_{\mathbf{M}_{T_0}}(\alpha, f_{T_0})}$ (by 19.6) $= \overline{\lambda T_0.V_{\mathbf{M}_{T_0}}(\alpha', f_{T_0})}$ (by the choice of U and the claim made in the above construction of $\mathbf{M}$) $= V_{\mathbf{M}}(\alpha', f)$. **QED.**

Instead of further exact definitions, a short presentation concludes this chapter, of one particular way of generalizing earlier concerns of correspondence and definability to this setting.

For any type $a = \langle b, c \rangle$, define the *subtypes* of a as b and c together with their subtypes. Now suppose that λ-abstraction occurs in a term α with respect to a variable of type a, but not with respect to variables whose type has a as a subtype. (a is "maximal with respect to λ-abstraction in α".) The definability questions of previous chapters may be posed here in the form "When is α equivalent to a term α' of the same type in which λ-abstraction only occurs with respect to types over which it occurred in α as well *but for a itself*?" (By "equivalent" it is meant that, for any $\mathbf{M}$ and f, $V_{\mathbf{M}}(\alpha, f) = V_{\mathbf{M}}(\alpha', f)$.) Of course, it has to be decided if "$\mathbf{M}$" ranges over all general models, or just over all standard models. In the first sense, the relation of equivalence is recursively enumerable; and results like the following (related to theorem 18.9) may be expected:

"α can be reduced to an α' as described if and only if, for all general models $\mathbf{M}$, $\mathbf{M}'$ differing only in their sets D_a (where a is the common type of α and α'), and the sets D_b affected by the difference as regards D_a, it holds that, for each suitable assignment f, $V_{\mathbf{M}}(\alpha, f) = V_{\mathbf{M}'}(\alpha, f)$".

In the second sense, reductions will become much more complicated; as before. Although L_ω is not a *ramified* theory of types, there would still seem to be a connection with Russel's "Axiom of Reducibility". For, the above question could also be interpreted as a search for the *syntactically simplest* definition for a given class of objects (of a certain type). Instead of merely postulating the existence of such definitions, as Russell did, one would like to study the question just when these exist at all.

BIBLIOGRAPHY

1. Van Benthem, J. F. A. K.: "A note on modal formulas and relational properties", «Journal of Symbolic Logic», vol. 40, 1975, pp. 85-88.
2. van Benthem, J. F. A. K.: "A problem on expansions and its connection with modal logic", unpublished paper, Instituut voor Grondslagenonderzoek, University of Amsterdam, 1975.
3. van Benthem, J. F. A. K.: "Canonical modal logics and ultrafilter extensions", «Journal of Symbolic Logic», vol. 44, 1979, pp. 1-8.
4. van Benthem, J. F. A. K.: "Four paradoxes", «Journal of Philosophical Logic», vol. 7, 1978, pp. 49-72.
5. van Benthem, J. F. A. K.: *Modal correspondence theory*, dissertation, Mathematisch Instituut & Instituut voor Grondslagenonderzoek, University of Amsterdam, 1976.
6. van Benthem, J. F. A. K.: "Modal formulas are either elementary or not $\Sigma\Delta$-elementary", «Journal of Symbolic Logic», vol. 41, 1976, pp. 436-438.
7. van Benthem, J. F. A. K.: "Modal logic as second-order logic", Report 77-04, Mathematisch Instituut & Instituut voor Grondslagenonderzoek, University of Amsterdam, 1977.
8. van Benthem, J. F. A. K.: "Modal reduction principles", «Journal of Symbolic Logic», vol. 41, 1976, pp. 301-312.
9. van Benthem, J. F. A. K.: "Ramsey eliminability", «Studia Logica», vol. 37, 1978, pp. 321-336.
10. van Benthem, J. F. A. K.: "Some correspondence results in modal logic", Report 74-05, Mathematisch Instituut & Instituut voor Grondslagenonderzoek, University of Amsterdam, 1974.
11. van Benthem, J. F. A. K.: "Some kinds of modal completeness", «Studia Logica», vol. 39, 1980, pp. 125-141.

12. van Benthem, J. F. A. K.: "Syntactic aspects of modal incompleteness theorems", «Theoria», vol. 45, 1979, pp. 63-77.

13. van Benthem, J. F. A. K.: "Two simple incomplete modal logics", «Theoria», vol. 44, 1978, pp. 25-37.

14. Blok, W. J.: *Varieties of interior algebras*, dissertation, Mathematisch Instituut, University of Amsterdam, 1976.

15. Bochenski, I. M.: *A history of formal logic*, University of Notre Dame Press, Notre Dame, 1961.

16. Bull, R. A.: "That all normal extensions of *S4.3* have the finite model property", «Zeitschrift für mathematische Logik und Grundlagen der Mathematik», vol. 12, 1966, pp. 341-344.

17. Chang, C. C. and Keisler, H. J.: *Model Theory*, North-Holland, Amsterdam, 1973.

18. Cresswell, M. J.: *Logics and languages*, Methuen, London, 1973.

19. Doets, H. C.: "Reduction of higher-order to second-order logic", unpublished paper, Instituut voor Grondslagenonderzoek, Amsterdam, 1976.

20. Enderton, H. B.: *A mathematical introduction to logic*, Academic Press, New York, 1972.

21. Feferman, S.: "Persistent and invariant formulas for outer extensions", «Compositio Mathematica», vol. 20, 1969, pp. 29-52.

22. Fine, K.: "An incomplete logic containing *S4*", «Theoria», vol. 40, 1974, pp. 23-28.

23. Fine, K.: "Logics containing *K4*. Part I", «Journal of Symbolic Logic», vol. 39, 1974, pp. 31-42.

24. Fine, K.: "Some connections between elementary and modal logic", *Proceedings of the third Scandinavian Logic Symposium, Uppsala 1973*, ed. by S. Kanger, North-Holland, Amsterdam, 1975, pp. 15-31.

25. Fine, K.: "The logics containing *S4.3*", «Zeitschrift für mathematische Logik und Grundlagen der Mathematik», vol. 17, 1971, pp. 371-376.

26. Fitch, F. B.: "A correlation between modal reduction principles and properties of relations", «Journal of Philosophical Logic», vol. 2, 1973, pp. 97-101.

27. Gabbay, D.: *Investigations into modal and tense logics, with applications to problems in linguistics and philosophy*, Reidel, Dordrecht, 1976.

28. Gödel, K.: "Eine Interpretation des intuitionistischen Aussagen-

kalküls", «Ergebnisse eines mathematischen Kolloquiums», vol. 4, 1933, pp. 34-40.

29. Goldblatt, R. I.: "First-order definability in modal logic", «Journal of Symbolic Logic», vol. 40, 1975, pp. 35-40.

30. Goldblatt, R. I.: "Metamathematics of modal logic", «Reports on Mathematical Logic», vol. 6, 1976, pp. 41-77, and vol. 7, 1976, pp. 21-52.

31. Goldblatt, R. I. and Thomason, S. K.: "Axiomatic classes in propositional modal logic", in J. Crossley, ed.: *Algebra and Logic*, Springer Lecture Notes in Mathematics 450, Berlin, 1974.

32. Grätzer, G.: *Universal Algebra*, Van Nostrand, Princeton, 1968.

33. Henkin, L.: "Completeness in the theory of types", «Journal of Symbolic Logic», vol. 15, 1950, pp. 81-91.

34. Hilpinen, R. (ed.): *Deontic Logic*, Reidel, Dordrecht, 1971.

35. Hintikka, J.: *Models for modalities*, Reidel, Dordrecht, 1969.

36. Hughes, G. E., and Cresswell, M. J.: *An introduction to modal logic*, Methuen, London, 1968.

37. Jech, Th. J.: *The axiom of choice*, North-Holland, Amsterdam, 1973.

38. Jennings, R. E., Schotch, P. K. and Johnston, D. K.: "Universal first-order definability in modal logic", «Zeitschrift für mathematische Logik und Grundlagen der Mathematik», vol. 26, 1980, pp. 327-330.

39. de Jongh, D. H. J. and Troelstra, A. S.: "On the connection of partially ordered sets with some pseudo-Boolean algebras", «Indagationes Mathematicae», vol. 28, 1966, pp. 317-329.

40. Kanger. S.: "The morning star paradox", «Theoria», vol. 23, 1957, pp. 1-11.

41. Kaplan, D.: Review of S. A. Kripke, "Semantical analysis of modal logic I", «Journal of Symbolic Logic», vol. 31, 1966, pp. 120-122.

42. Krabbe, E. C. W.: "Propositionele Tijdslogika", master's thesis, Instituut voor Grondslagenonderzoek, Amsterdam, 1972.

43. Kripke, S.: "A completeness theorem in modal logic", «Journal of Symbolic Logic», vol. 24, 1959, pp. 1-14.

44. Kripke, S.: "Naming and necessity", in: *Semantics of natural language*, ed. by D. Davidson and G. Harman, Reidel, Dordrecht, 1972.

45. Kripke, S.: "Semantic analysis of modal logic. I", «Zeitschrift für mathematische Logik und Grundlagen der Mathematik», vol. 9, 1963, pp. 67-96.

46. Kripke, S.: "Semantical considerations on modal logic", «Acta Philosophica Fennica», vol. 16, 1963, pp. 83-94.

47. Lachlan, A. H.: "A note on Thomason's refined structures for tense logic", «Theoria», vol. 40, 1974, pp. 117-120.

48. Lemmon, E. J. and Scott, D.: *An introduction to modal logic*, (formerly called "Intensional logic"), ed. by K. Segerberg, Blackwell, Oxford, 1977.

49. Lewis, C. I. and Langford, C. H.: *Symbolic logic*, Dover Reprint, New York, 1959.

50. Lewis, D.: "Counterpart theory and quantified modal logic", «Journal of Philosophy», vol. 65, 1968, pp. 113-126.

51. Lewis, D.: "Intensional logics without iterative axioms", «Journal of Philosophical Logic», vol. 3, 1974, pp. 457-466.

52. Lindström, P.: "On extensions of elementary logic", «Theoria», vol. 35, 1969, pp. 1-11.

53. Löb, M. H.: "Solution to a problem of Leon Henkin", «Journal of Symbolic Logic», vol. 20, 1955, pp. 115-118.

54. Magari, R.: "Representation and duality theory for diagonalizable algebras", «Studia Logica», vol. 34, 1975, pp. 305-313.

55. Makinson, D. C.: "A normal calculus between *T* and *S4* without the finite model property", «Journal of Symbolic Logic», vol. 34, 1969, pp. 35-38.

56. Makinson, D. C.: "On some completeness theorems in modal logic", «Zeitschrift für mathematische Logik und Grundlagen der Mathematik», vol. 12, 1966, pp. 379-384.

57. Makinson, D. C.: "Some embedding theorems for modal logic", «Notre Dame Journal of Formal Logic», vol. 12, 1971, pp. 252-254.

58. McKinsey, J. C. C. and Tarski, A.: "Some theorems about the sentential calculi of Lewis and Heyting", «Journal of Symbolic Logic», vol. 13, 1948, pp. 1-15.

59. Montague, R.: *Formal Philosophy*, Yale University Press, New Haven, 1974.

60. Montague, R.: "Syntactical treatments of modality, with corollaries on reflexion principles and finite axiomatizability", «Acta Philosophica Fennica», vol. 16, 1963, pp. 153-167.

61. Mortimer, M.: "Some results in modal model theory", «Journal of Symbolic Logic», vol. 39, 1974, pp. 496-508.

62. Orey, S.: "Model theory for the higher-order predicate calculus", «Transactions of the American Mathematical Society», vol. 92, 1959, pp. 72-84.

63. Plantinga, A.: *The nature of necessity*, Clarendon, Oxford, 1974.

64. Prior, A. N.: *Past, present and future*, Oxford University Press, Oxford, 1967.

65. Quine, W. V. O.: "Reference and modality", in: *From a logical point of view*, Harper and Row, New York, 1961.

66. Sahlqvist, H.: "Completeness and correspondence in the first and second-order semantics for modal logic", *Proceedings of the third Scandinavian Logic Symposium, Uppsala 1973*, ed. by S. Kanger, North-Holland, Amsterdam, 1975, pp. 110-143.

67. Segerberg, K.: *An essay in classical modal logic*, Filosofiska Studier 13, Uppsala, 1971.

68. Segerberg, K.: "Decidability of *S4.1*", «Theoria», vol. 34, 1968, pp. 7-20.

69. Segerberg, K.: "Modal logics with linear alternative relations", «Theoria», vol. 36, 1970, pp. 301-322.

70. Segerberg, K.: "Post completeness in propositional modal logic", «Journal of Symbolic Logic», vol. 37, 1972, pp. 711-715.

71. Segerberg, K.: Review of M. K. Rennie, "On postulates for temporal order", «Journal of Symbolic Logic», vol. 37, 1972, p. 629.

72. Shoenfield, J. R.: *Mathematical logic*, Addison-Wesley, Reading, 1967.

73. Smoryński, C.: "Investigations of intuitionistic formal systems by means of Kripke models", dissertation, Stanford University, 1973.

74. Solovay, R. M.: Provability interpretations of modal logic", «Israel Journal of Mathematics», vol. 25, 1976, pp. 287-304.

75. Thomason, S. K.: "An incompleteness theorem in modal logic", «Theoria», vol. 40, 1974, pp. 30-34.

76. Thomason, S. K.: "Non-compactness in propositional modal logic", «Journal of Symbolic Logic», vol. 37, 1972, pp. 716-720.

77. Thomason, S. K.: "Reduction of second-order logic to modal logic", «Zeitschrift für mathematische Logik und Grundlagen der Mathematik», vol. 21, 1975, pp. 107-114.

78. Thomason, S. K.: "Reduction of tense logic to modal logic I", «Journal of Symbolic Logic», vol. 39, 1974, pp. 549-551.

79. Thomason, S. K.: "Semantic analysis of tense logics", «Journal of Symbolic Logic», vol. 37, 1972, pp. 150-158.

Finito di stampare
nel novembre 1985
dalla Grafitalia s.r.l.
Cercola (Napoli)