JOHAN VAN BENTHEM

CORRESPONDENCE THEORY

1 INTRODUCTION TO THE SUBJECT

Correspondences

When possible worlds semantics arrived around 1960, one of its most charming features was the discovery of simple connections between existing intensional axioms and ordinary properties of the alternative relation among worlds. Decades of syntactic labour had produced a jungle of intensional axiomatic theories, for which a perspicuous semantic setting now became available. For instance, typical completeness theorems appeared such as the following:

> A modal formula is a theorem of **S4** if and only if it is true in all *reflexive, transitive* Kripke frames.

Indeed, **S4** may also be shown to be the modal logic of the *partial orders*; which matches the most famous modal logic with perhaps the most basic type of classical relational structure. Such matchings extend to logics higher up in the **S4**-spectrum. For instance, **S4.2** with its additional axiom

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

is complete with respect to those frames which are reflexive, transitive and *directed*, or *confluent*:

$$\forall xyz((Rxy \wedge Rxz) \to \exists u(Ryu \wedge Rzu))$$

Again, the latter condition is a 'diamond property' of classical fame.

Completeness results such as these have inspired a flourishing area of intensional *Completeness Theory*, witness the classic [Segerberg, 1971]. It took modal logicians some time, however, to realise that there are also direct semantic equivalences involved here, having nothing to do with deduction in modal logics. Indeed, the whole present *Correspondence Theory* arose out of simple observations such as the following, made in the early seventies.

EXAMPLE 1. The T-axiom $\Box p \to p$ is true in a Kripke frame $\langle W, R\rangle$ if and only if R is reflexive.

Here, 'true in a frame' means true in all worlds, under all assignments to the proposition letters.

Proof. '$\Rightarrow$': Consider any $w \in W$. If $\Box p \to p$ is true in $\langle W, R\rangle$, then, in particular, it is true at w under the assignment V with

$$V(p) = \{v \in W \mid Rwv\}.$$

Thus, $\Box p$ will be at w true by definition — and, hence, also p: i.e. Rww.

'$\Leftarrow$': By reflexivity, truth at all R-alternatives implies actual truth. ■

EXAMPLE 2. The **S4**-axiom $\Box p \to \Box\Box p$ is equivalent to transitivity.

Proof. By an analogous argument. ■

EXAMPLE 3. The **S4.2**-axiom $\Diamond\Box p \to \Box\Diamond p$ defines directedness.

Proof. '$\Rightarrow$': Consider arbitrary $w, v, u \in W$ such that Rwv, Rwu. Let the assignment V have

$$V(p) = \{s \in W \mid Rvs\}.$$

Immediately, this gives truth of $\Box p$ at v. Therefore, $\Diamond\Box p$ is true at w, whence $\Box\Diamond p$ must hold as well. It follows that $\Diamond p$ is true at u; i.e. u has some R-successor in $V(p)$ — whence v, u share a common R-successor.

'$\Leftarrow$': If $\Diamond\Box p$ is true at W, say because of some v with Rwv verifying $\Box p$, then $\Diamond p$ will be true at all R-successors of w. For, all of these share at least one successor with v, by directedness. ■

Not all correspondences are equally simple. For instance, **S4.2** has a companion logic **S4.1** obtained by enriching **S4** with the 'McKinsey Axiom' $\Box\Diamond p \to \Diamond\Box p$. This converse of the **S4.2** axiom turns out to be much more complex. A well-known completeness theorem says that **S4.1** axiomatises the modal theory of those Kripke frames which are reflexive, transitive as well as *atomic*:

$$\forall x \exists y (Rxy \wedge \forall z (Ryz \to z = y)).$$

(Notice that we need *identity* here, in addition to the predicate constant R.) We shall see later in Section 2.2 that the **S4.1** axioms together (just) manage to define the above threefold relational condition, but that the McKinsey Axiom does not define atomicity on its own (it is weaker). Indeed, this simple modal principle does not possess a first-order relational equivalent at all — a discovery made independently by several people around 1975.

Modal Formulas as Conditions on the Alternative Relation

The general picture emerging here is that of modal axioms expressing certain 'classical' constraints on the alternative relation in frames where they are valid. With hindsight, this observation is hardly surprising. After all, given some valuation, the clauses of the basic Kripke truth definition amount to a *translation* from modal formulas into classical ones involving R. Thus, e.g.,

$$\begin{array}{lll} \Box p \to p & \text{becomes} & \forall y(Rxy \to Py) \to Px \\ \Box p \to \Box\Box p & \text{becomes} & \forall y(Rxy \to Py) \to \\ & & \quad \to \forall y(Rxy \to \forall z(Ryz \to Pz)), \end{array}$$

while the McKinsey Axiom $\Box\Diamond p \to \Diamond\Box p$ becomes

$$\forall y(Rxy \to \exists z(Ryz \wedge Pz)) \to \exists y(Rxy \wedge \forall z(Ryz \to Pz)).$$

Here the parameter 'x' refers to the current world of evaluation, while unary predicate constants P $(Q, \ldots)$ denote the sets of worlds where the corresponding proposition letter p $(q, \ldots)$ holds.

Let us pause, to realise how, by this simple observation alone, many established results about classical predicate logic can be transferred straightaway to modal logic. For instance, for Kripke frames plus a fixed assignment (the modal 'models' of Section 2.1), *Compactness* and *Löwenheim–Skolem* results are immediate. If, e.g. a set of modal formulas is finitely satisfiable in Kripke models (given suitable assignments), then its classical transcription will be finitely satisfied too. Hence, by ordinary compactness, the latter set is simultaneously satisfied in some structure $\langle W, R; P, Q, \ldots\rangle$: which forms a Kripke frame cum assignment verifying the original set.

But, this perspective is not quite the one we need.

In the evaluation of modal formulas according to the above truth definition, two factors are intermingled: the relational pattern of the worlds and the particular 'facts', i.e. the assignment. But the latter — the particular denotations of constants $P, Q, \ldots$ — is not relevant to the role of modal formulas as relational constraints. Indeed, these may even obscure the issue. When, e.g. $V(p)$ equals W, $\Box p \to p$ holds in all worlds — but this observation is completely uninformative about the true content of this axiom (viz. reflexivity).

In order to arrive at the proper perspective, one simply abstracts from the effects of particular assignments, by means of a *universal* quantification over the unary predicates in the preceding translation. Thus, for instance,

$$\Box(p \vee q) \to (\Box p \vee \Box q)$$

now becomes

$$\begin{array}{l} \forall P \forall Q\ (\forall y(Rxy \to (Py \vee Qy)) \to (\forall y(Rxy \to Py) \vee \\ \quad \vee \forall y(Rxy \to Qy))). \end{array}$$

Notice that modal formulas now get *second-order* transcriptions, as opposed to the earlier first-order ones.

The parameter 'x' has remained: the present relational conditions are still 'local' in some actual world. A 'global' condition is obtained by performing one more universal quantification, this time with respect to this world parameter. The distinction is not without importance. The local version is more suitable for the original Kripke structures $\langle W, R, w_0 \rangle$, in which some 'actual world' w_0 figured prominently, as well as for 'non- normal' modal semantics, in which certain worlds are distinguished from others. The global reading is the more common one, however, which will be predominant in the sequel.

Again, the very point of view embodied in the above translation is significant — even though some of the earlier transfer phenomena are lost. What is lost, for instance, are most useful forms of compactness, as well as the Löwenheim–Skolem property. There is no automatic guarantee through second-order logic that, if a modal formula is true in some uncountable Kripke frame (i.e. under *all* valuations) it will be true in its countable elementary subframes (again, under all valuations). Still, this very phenomenon will be used to drive a wedge between 'essentially first-order' and 'essentially second-order' modal axioms in Section 2.2. Moreover, not all is lost. The above transcriptions are very simple second-order formulas, viz. so-called Π^1_1-*sentences*, with all second-order quantifiers occurring in a universal prefix in front of a first-order matrix. From classical logic, we still now a few things about Π^1_1-sentences, that will turn out useful. (Cf. the chapters on Higher Order Logic and Algorithms in Volume 1 of this *Handbook* for background.)

One such thing is involved in the following obvious question. In the light of earlier examples of correspondence, the present second-order transcriptions are exceedingly cumbersome. Compare, e.g. for the T-axiom $\Box p \to p$,

$$\forall x Rxx \text{ with } \forall x \forall P(\forall y(Rxy \to Py) \to Px).$$

Yet it was the discovery of the former simple *first-order* equivalents that motivated the above investigation in the first place. Now for some modal formulas, the second-order complexity may be unavoidable — witness the example of McKinsey's Axiom. But at least, there arises an obvious basic

QUERY: Which modal formulas define first- order relational conditions — and how do they manage it?

By the above perspective, classical sources provide one immediate answer. A Π^1_1-sentence is first-order definable if and only if it is preserved under the formation of *ultraproducts*, a fundamental construction in classical model theory. Through the above transcription, the same criterion applies to modal formulas. (The technical ins and outs of this point, as well as of related ones in this introduction, are postponed until the relevant sections: Sections 2.1 and 2.2 in this case.)

Modal Correspondence Theory

The preceding query has been the starting point for a systematic study of classical definability of modal formulas, when viewed as relational principles. Now the mentioned ultraproduct characterisation is a very abstract, global one, rather removed from the actual business of finding correspondences. Also historically, it is a rather late development — and we shall therefore turn to more concrete themes, as they evolved.

At first sight, *proving* first-order definability seems a simple matter: just find an equivalent, and show that it works. Still, there is the question how much system there is to this activity. For instance, Examples 1–3 exhibited regularities in their proofs. And indeed, closer inspection reveals that reflexivity, transitivity and directedness may be obtained from the second-order transcriptions of the **S4.2**-axioms through certain *substitutions* of 'minimal' *definable assignments.*

The heuristics behind this method is simply this. If, e.g. $\Box p \to p$ is true at x, then the most 'parsimonious' way of verifying the antecedent (i.e. by having $V(p) = \{y \mid Rxy\}$) carries maximal information about the whole implication. This essentially, is why the substitution of Rxu for Pu in

$$\forall x \forall P(\forall y(Rxy \to Py) \to Px)$$

yields the equivalent formula

$$\forall x(\forall y(Rxy \to Rxy) \to Rxx).$$

By the universal validity of the antecedent, the latter may be simplified to the usual statement of reflexivity. A completely analogous line of thought produces transitivity from the transcription of $\Box p \to \Box\Box p$. Some complications arise with antecedents as in $\Diamond\Box p \to \Box\Diamond p$; but the general idea remains the same. In this way, one discovers a large recursive class of modal formulas with effectively obtainable first-order equivalents.

Nevertheless, this method of substitutions also has definite limits. Notably, it does not work for all first-order definable modal formulas — as will be proved in Section 2.2 for the case of **S4.1**. In connection with this matter, the exact *combinatorial complexity* of the set of first-order definable modal formulas is still unknown — but there are reasons for fearing that it is not even arithmetically definable (let alone, recursive or recursively enumerable).

Disproving first-order definability is a more difficult matter. Indeed, how should one go about this at all? The common pattern in all examples in the literature comes to this: find some semantic preservation property of first-order sentences, which is lacked by the modal formula under consideration. Thus, e.g. the earliest published contribution by the present author was an example showing how the McKinsey Axiom sins against the Löwenheim–

Skolem theorem. It holds in a certain uncountable Kripke frame (to be presented in Section 2.2.) without holding in any of a certain group of its countable elementary subframes. A classical example of this phenomenon occurs when Dedekind Continuity (itself a Π_1^1-property) is added to the first-order ordering theory of the rationals. The resulting Π_1^1-sentence has uncountable models (notably, the reals); but, it even lacks countable models altogether.

The modal examples of 'essentially second-order' axioms to be found in Section 2.2 will serve to delimit the range of the above method of substitutions. As so often, the McKinsey Axiom again provides an illuminating example. The above heuristics of 'minimal verification' typically fails for antecedents such as $\Box\Diamond p$, expressing some dependency — and first-order failure is immediate.

Besides the modal half of the story, so to speak, there also exists the opposite direction, looking from classical formulas to modal ones. Again, this inspires a basic

QUERY. Which first-order relational conditions are modally definable?

The 'positive' side of this matter again concerns the establishing of valid equivalences. Thus, for instance, how does one find a modal definition for such a classical favourite as *connectedness*

$$\forall xyz((Rxy \wedge Rxz) \rightarrow (Ryz \vee Rzy)))?$$

This time, the heuristics consists in imagining a situation where the property fails, together with a way of 'maximally exploiting' this failure through modal formulas. In the above particular case, supposing that $Rxy, Rxz, \neg Ryz, \neg Rzy$, one sets $\Box p$ true at y (with p false at z) and $\Box q$ true at z (with q false at y). This has the effect of verifying the following formula at x:

$$\Diamond(\Box p \wedge \neg q) \wedge \Diamond(\Box q \wedge \neg p).$$

Now, the original property itself will correspond to the negation of this modal 'failure description', i.e.

$$\neg(\Diamond(\Box p \wedge \neg q) \wedge \Diamond(\Box q \wedge \neg p)).$$

By some familiar equivalence transformations, this becomes

$$\Box(\Box p \rightarrow q) \vee \Box(\Box q \rightarrow p),$$

a principle known from the literature as Geach's Axiom.

It remains to be shown, of course, that conversely, failure of this axiom implies failure of connectedness; but this is immediate. In order to cross-check, one might also apply the earlier method of substitutions to (some suitable transform of) the Geach Axiom: and indeed, connectedness will ensue.

The 'negative' side again consists of disproofs. Here as well, these turn out to possess a particular interest — as we are forced to contemplate 'typical behaviour' of modal formulas. A standard example is the following. Although reflexivity was modally definable, *irreflexivity* turns out intractable: $\forall x \neg Rxx$. But, failed attempts are no definite refutations. What we need is some semantic property of modal formulas, as relational conditions on Kripke frames, which is not shared by this particular first-order sentence.

At this point, the modal model theory of Section 2.1 comes in. There, one finds that the following mappings play a fundamental role in the transmission of modal truth between Kripke frames: a *p-morphism* is a function f from a frame $\langle W, R_1 \rangle$ to a frame $\langle W_2, R_2 \rangle$ which

1. preserves R_1, and
2. 'almost' preserves R_2, in the following sense:
   'If $R_2 f(w)v$, then there exists some $u \in W_1$ such that (a) $R_1 wu$ and (b) $f(u) = v$'.

Under different names, this notion has had a career in standard logic already, e.g. the 'Mostowski collapse' in set theory is of this kind.

For the purposes of the present example, it need only be recorded that subjective p-morphisms preserve truth of modal formulas on Kripke frames. But then, irreflexivity may be dismissed: it holds in the frame of the natural numbers with the usual order, but it fails in its p-morphic image (!) arising from the contraction to one single reflexive point.

This example will have given a taste of the actual field-work in this area of Correspondence Theory. There also arises the more general question, of course, whether some combination of modally valid preservation requirements manages to *characterise* all and only the modally definable first-order sentences. This is indeed the case, and an elegant result to this effect — involving p-morphisms as well as other basic constructions, will be proved in Section 2.4.

The preceding survey by no means exhausts the range of questions that can be investigated in Correspondence Theory — but it does convey the spirit.

Correspondence and Completeness

Three pillars of wisdom support the edifice of Modal Logic. There is the ubiquitous *Completeness Theory*, the present Correspondence, or, more generally, *Definability Theory* — and finally, the *Duality Theory* between Kripke frames and 'modal algebras' (cf. Section 2.3 below) has become an area of its own. Connections between the latter two will become apparent as Section

2 unfolds — in particular, the above-mentioned characterisation of modally definable first-order sentences will be obtained as a consequence of the classic Birkhoff Theorem of Universal Algebra, applied to modal algebra.

The relation between correspondence and completeness is less vital to subsequent developments. Moreover, it turns out to be rather complex — and indeed, only partially understood. Nevertheless, for those readers who are familiar with the basic notions of Completeness Theory, the following sketch of issues may serve to bring questions of correspondence closer to traditional concerns.

The early completeness theorems in modal logic were brought under one heading in [Segerberg, 1971]: 'modal logic **L** is determined by a class $\mathfrak{K}$ of Kripke frames', i.e. **L** axiomatises the modal theory of $\mathfrak{K}$ (on the basis of the minimal logic **K**).

As before, two perspectives emerge here. First, one may start with a given class $\mathfrak{K}$, asking for a recursive axiomatisation **L** of its modal theory. In general, there is no guarantee for success here; but there is one helpful observation involving first-order definability.

FACT 4. If $\mathfrak{K}$ is elementary (i.e. defined by a single first-order sentence), then its modal theory is recursively axiomatisable.

Proof. Let $\alpha = \alpha(R, =)$ define $\mathfrak{K}$. A modal formula φ belongs to the theory of $\mathfrak{K}$ if and only if it holds in all frames in $\mathfrak{K}$. This may be restated as follows:

$$\alpha \vDash \forall x \forall P_1 \ldots \forall P_n \tau(\varphi);$$

where $\tau(\varphi)$ is the earlier first-order translation of φ, while $p_1, \ldots, p_n$ are the proposition letters occurring in the latter formula. Now, the predicate variables $P_1, \ldots, P_n$ do not occur in the first-order sentence α, and, therefore the above implication is equivalent to $\alpha \vDash \forall x \tau(\varphi)$. But this is an ordinary first-order implication. So, since the latter notion is recursively axiomatisable, the same must be true for membership of the modal theory of $\mathfrak{K}$.

Axiomatisable, yes, but axiomatisable on the basis of the minimal modal logic **K**? Even this is true, choosing a suitable recursive set of axioms as in the proof of Craig's Theorem in classical logic and noticing that **K** contains *modus ponens* (which is all that is needed). ■

Thus, in retrospect, the earlier completeness theorems for reflexive, transitive orders (and other elementary classes) were quite predictable.

The direction from classes of frames to logics is not the current one in modal logic; being more appropriate to areas such as tense logic, where temporal structures often precede temporal theories. Usually, one already possesses a certain logic **L**, asking for a class $\mathfrak{K}$ of Kripke frames with respect to which it is complete. (Notice that, if *any* class $\mathfrak{K}$ suffices, then the whole class of Kripke frames validating **L** will.)

Nowadays, we know that not all modal logics are in fact *complete* in the above sense, contrary to earlier expectations. This is the content of the celebrated 'modal incompleteness theorems' in [Fine, 1974; Thomason, 1974]. But it has been hoped that, at least, all *first-order definable* axiom sets are complete. (Indeed, a defective proof to this effect has circulated.) Even this more modest expectation was frustrated in [van Benthem, 1978]:

FACT 5. The modal logic **L** with characteristic axioms

$$\Box p \to p$$
$$\Box\Diamond p \to \Diamond\Box p$$
$$(\Diamond p \wedge \Box(p \to \Box p)) \to p$$

is first-order definable: its frames are just those satisfying the condition

$$\forall xy(Rxy \leftrightarrow x = y).$$

But the characteristic axiom of the modal theory of the latter class of frames, viz. $\Box p \leftrightarrow p$, is not minimally derivable from **L**.

The relevant correspondence will be proved in Section 2.2. For the moment, it may be noticed that the third axiom defines a notion of 'safe return': from any R-successor of a world x, one can always return to x by following some finite R-chain of R-successors of x.

The relevant argument is highly nontrivial, far outside the range of our earlier method of substitutions. Nevertheless, even the latter has its relevance for completeness theory, as we shall see presently.

What the modal incompleteness theorems show is that the minimal modal logic **K** is to weak to produce all modally valid inferences. But of course, there may be stronger reasonable 'base logics'. One particular example arises from the method of substitutions. For instance, in proving the equivalence of substitution instances with more current first-order conditions, one uses an extremely natural second-order logic $\mathbf{K}_2$ with the following deductive apparatus:

> Some first-order base complete with respect to *modus ponens*,
> similar axioms for the second-order quantifiers;

with the following form of 'first-order instantiation' allowed for first-order formulas ψ

$$\forall x\varphi(X) \to \varphi(\psi).$$

Through the earlier second-order transcription, $\mathbf{K}_2$ may be used as a modal base logic.

Here is an example of some fame. In the metamathematics of arithmetical provability (cf. [Boolos, 1979] or Smoryński's in a later volume of this *Handbook*), the following two modal axioms are basic:

$$\Box p \to \Box\Box p, \quad \Box(\Box p \to p) \to \Box p \quad \text{('Löb's Axiom')}.$$

The semantic import of the latter will be established in Section 2.2: it holds in those Kripke frames whose alternative relation is transitive, while possessing a well-founded converse. Moreover, transitivity is $\mathbf{K}_2$-derivable from Löb's Axiom, by the substitution of

$$Rxu \wedge \forall y(Ruy \rightarrow Rxy) \quad \text{for } Pu.$$

(The antecedent becomes universally valid, while the consequent expresses transitivity.) An advantage of $\mathbf{K}_2$ over $\mathbf{K}$? No, around 1975, Dick de Jongh and Giovanni Sambin found a K-deduction for the first axiom from the second after all. The two deductions are related, but systematic connections between $\mathbf{K}$-deductions and $\mathbf{K}_2$-deductions have not been explored up to date.

Nevertheless, $\mathbf{K}_2$ is non-conservative over $\mathbf{K}$ in the modal realm. In [van Benthem, 1979b] we find the following incompleteness theorem.

FACT 6. The modal axiom

$$\Diamond\Box\bot \vee \Box(\Box(\Box p \rightarrow p) \rightarrow p),$$

with $\bot$ the falsum, defines the same class of Kripke frames as $\Diamond\Box\bot \vee \Box\bot$. But, the latter formula is not $\mathbf{K}$-derivable from the former — even though it is $\mathbf{K}_2$-derivable.

Again, there is a correspondence involved here. But the idea is illustrated by a simple $\mathbf{K}_2$-deduction at the back of this result:

1. $\forall P(\forall y(Rxy \rightarrow (\forall z(Ryz \rightarrow Pz) \rightarrow Py)) \rightarrow Px)$ $\quad ('\Box(\Box p \rightarrow p) \rightarrow p')$,
2. $\forall y(Rxy \rightarrow (\forall z(Ryz \rightarrow z \neq x) \rightarrow y \neq x)) \rightarrow x \neq x$ $\quad (x \neq u$ for $Pu)$,
3. $\neg\forall y(Rxy \rightarrow (\forall z(Ryz \rightarrow z \neq x) \rightarrow y \neq x))$,
4. $\exists y(Rxy \wedge \forall z(Ryz \rightarrow z \neq x) \wedge y = x)$,
5. $Rxx \wedge \forall z(Rxz \rightarrow z \neq x)$
6. $x \neq x$: a contradiction ($\bot$).

That $\mathbf{K}_2$, in its turn, must be modally incomplete (as is any proposed recursively axiomatised base logic) follows from the *general* incompleteness results in [Thomason, 1975].

First-order definability does not imply completeness. But, *when* a modal logic is both first-order definable and complete, it enjoys a very pleasant form of the latter property — viz. with respect to the underlying frame of its own *Henkin model*. ('First-order definability plus completeness imply canonicity': cf. [Fine, 1975; van Benthem, 1980].) Such *canonical* modal logics will be characterised semantically in Section 2.4: notice that many of the familiar text book examples are of this kind. In fact, a canonical completeness proof, such as that for **S4**, often proceeds by means of first-order conditions on the Henkin model, induced by the corresponding axioms.

The relation between these familiar 'Henkin arguments' and the above method of substitutions is at present still rather mysterious. Sahlqvist [1975] contains many examples of parallels; but Fine [1975] presents a problem. The modal formula

$$\Diamond\Box(p \vee q) \to \Diamond(\Box p \vee \Box q)$$

axiomatises a canonical modal logic, without being first-order definable. Thus, we are still far from complete clarity in the area between completeness and correspondence.

Variations and Generalisations

Logical model theory may be viewed as a marriage between ontology and language (or 'mathematics' and 'linguistics'). Accordingly, the semantics of propositional modal logic, our paradigm example up till now, exhibits the familiar triangle

$$\text{language} \xrightarrow[\text{interpretation}]{} \text{structures}$$

Or, from the above translational point of view, the components are

$$\text{prima facie language} \xrightarrow[\text{translation}]{} \text{representation language}$$

All these 'degrees of freedom' may be varied in intensional logic — and thus there appears a whole family of 'correspondence theories'. We shall explore some examples of recognised importance in Section 3. Here, let us just think about the various possibilities and their implications.

Even within the domain of propositional modal logic, alternatives have been proposed for Kripke-type relational semantics. Jennings, Johnstone and Schotch [1980] contains the proposal to work with *ternary* alternative relations, employing the following notion of necessity:

$$\Box\varphi \text{ is true at } x \text{ if } \forall yz(Rxyz \to \varphi(y) \vee \varphi(z)).$$

Their motivation was, amongst others, to create room for 'non-cumulation' of necessities: the 'Aggregation Axiom'

$$\Box p \wedge \Box q \to \Box(p \wedge q)$$

will no longer be valid. What happens to earlier correspondences in this new light? Old boundaries start shifting; e.g. $\Box p \to p$ remains first-order definable, but $\Box p \to \Box\Box p$ becomes essentially second-order on this semantics.

This is compensated for by the phenomenon of formerly unexciting principles, such as the Aggregation Axiom (which was trivially valid before) springing into unexpected bloom:

EXAMPLE 7. $\Box p \wedge \Box q \to \Box(p \wedge q)$ defines

$$\forall xyz(Rxyz \to (y = z \vee Rxyy \vee Rxzz)).$$

Proof. '$\Rightarrow$': Suppose the condition fails at x, y, z. Setting

$$V(p) = W = \{z\}, V(q) = W - \{y\},$$

will then verify $\Box p, \Box q$ at x, while $\Box(p \wedge q)$ is falsified (by $Rxyz$).

'$\Leftarrow$': Suppose that $\Box p, \Box q$ hold at x, and consider $Rxyz$. Either $y = z$, whence y verifies both p and q (by $Rxyy$ and the truth definition), or $Rxyy$, implying the same conclusion, or $Rxzz$, in which case z verifies both p and q. So, $\Box(p \wedge q)$ holds at x. ∎

As for the general theorems, forming the backbone of the subject, nothing essential changes in this ternary semantics.

This example changed both the structures and the form of the truth definition. What may not be generally realised is the variety offered even when fixing the two parameters of 'language' and 'structures'. Therefore, a short digression is undertaken here.

The Kripke truth definition is not sacrosanct — other clauses would have been quite imaginable. Thus, for instance, we may make the following

OBSERVATION 8. The truth definition '$\Box\varphi$ is true at x if $\forall y((Rxy \vee Ryx) \to \varphi(y))$' yields as a modal base logic **KB**; i.e. the minimal logic **K** plus the Brouwer Axiom $p \to \Box\Diamond p$.

Proof. The Brouwer Axiom defines *symmetry* of the alternative relation; as may be seen by substituting $u = x$ for Pu. And indeed **KB** is complete with respect to the class of symmetric Kripke frames. Hence, any non-theorem φ of **KB** is falsified on some symmetric frame $\langle W, R\rangle$. But, on symmetric frames R coincides with the relation $\lambda xy.\ (Rxy \vee Ryx)$ (i.e. R united with its converse $\breve{R}$); whence φ also fails by the new evaluation.

Conversely, if φ has a counter-example $\langle W, R\rangle$ under the new truth definition, then it has $\langle W, R \cup \breve{R}\rangle$ for an ordinary symmetric counter-example; whence it is outside of **KB**. ∎

Thus, there is a possible *trade-off* between truth definition and requirements on the alternative relation. The exact extent of this phenomenon remains to be investigated. Notice for example how **KB** is equally well generated by the following truth definition:

$$\Box\varphi \text{ is true at } x \text{ if } \forall y((Rxy \wedge Ryx) \to \varphi(y)).$$

The general principle behind such examples is this.

FACT 9. If $C(R)$ is any condition on R, and $\gamma(x, y)$ some formula in $R, =$ such that

1. If $C(R)$ is satisfied, then R and $\lambda xy.\gamma(x, y)$ coincide,
2. $\lambda xy.\gamma(x, y)$ satisfies C,

then the modal logic determined by (the Kripke frames obeying) C may also be generated without conditions through the truth definition

$$\Box\varphi \text{ is true at } x \text{ if } \forall y(\gamma(x, y) \rightarrow \varphi(y)).$$

This rather subversive shift in perspective will not be investigated in this contribution. At this point, it merely serves to remind us that not a single aspect of the semantic enterprise is immune to revision.

Leaving the realm of modal logic, of the many intensional candidates for a correspondence perspective, only a few have been explored up to date. In Section 3, some important examples are reviewed briefly, viz. *tense logic, conditional logic* and *intuitionistic logic*. These illustrate, in ascending order, certain difficulties which tend to make Correspondence Theory rather more difficult (often also: more exciting) in many cases. These difficulties have to do with 'pre-conditions' on the alternative relation (not very serious), and the phenomenon of 'admissible assignments' (rather more serious), to be explained in due course. Nevertheless, for instance, Intuitionistic Correspondence Theory will turn out to possess also some elegant features lacked by its modal predecessor.

A few examples, even without proof, will render the above remarks more concrete. In tense logic, the correspondence runs between temporal axioms and properties of the temporal order ('before', 'earlier than').

EXAMPLE 10 ('Hamblin's Axiom'). $(p \wedge Hp) \rightarrow FHp$ defines discreteness of Time:

$$\forall x \exists y > x \forall z < y \; (z = x \vee z < x).$$

In the logic of counterfactual conditionals, conditional inferences are related to the behaviour of the comparative similar ordering C among alternative worlds.

EXAMPLE 11 (Stalnaker's Axiom of 'Conditional Excluded Middle'). $(p \Rightarrow q) \vee (p \Rightarrow \neg q)$ defines linearity of alternative worlds:

$$\forall xyz(y = z \vee Cxyz \vee Cxzy).$$

Finally, in intuitionistic logic, ('intermediate') axioms impose constraints upon the possible growth patterns of stages of knowledge.

EXAMPLE 12 ('Weak Excluded Middle'). $\neg p \vee \neg\neg p$ defines 'local convergence' of growing stages, i.e. directedness:

$$\forall xyz((x \subseteq y \wedge x \subseteq z) \rightarrow \exists u(y \subseteq u \wedge z \subseteq u)).$$

Proofs, and further explorations are postponed until the relevant sections. At this stage, the experienced reader may predict that two nuts will be especially difficult to crack for any Correspondence Theory.

The first of these concerns the earlier tacit restriction to *propositional* logic: what happens in the predicate case? In Section 2.5 we shall see that no essential problems seem to arise — although the field remains largely unexplored.

A more formidable problem arises when the truth definition for the intensional operators itself becomes of higher-order complexity. In that case, e.g. a search for possible first-order equivalents of intensional axioms seems rather pointless. This eventuality arises when disjunction is evaluated barwise in Beth semantics for intuitionistic logic (i.e. $\varphi \vee \psi$ is true at x if the φ-worlds and ψ-worlds together form a barrier intersecting each branch passing through x).

The last word has not been said here, however. Philosophically, it seems a rather unsatisfactory division of semantic labour to let the truth definition absorb structural complexity (in this case: the second-order behaviour of branches). The latter should be located where it belongs, viz. in the structures themselves. And indeed, the Beth semantics admits of a two-sorted first-order reformulation in terms of nodes and paths, which generates a Correspondence Theory of the usual kind.

All this is not to say that there are no limits to the useful application of a correspondence perspective. But, these are to be found in *philosophical* relevance, rather than technical *impossibility.* One should study correspondences only as long as they serve the purpose of semantic enlightenment — which is the shedding of light upon one conceptual framework by relating it systematically to another.

2 MODALITY

In this chapter, modal correspondence theory will be surveyed against the background of modal model theory and modal algebra, whose basics are explained. (Cf. the chapter by Bull and Segerberg in this volume for the necessary background.)

2.1 Modal Model Theory

The basic structures of modal semantics are introduced: *frames, models* and *general frames.* These may be studied either purely classically, or with a specifically modal purpose. In both cases, the emphasis is not upon such structures in isolation, but upon their 'categorial context': what are their relations with other structures, and which of these relations are truth-preserving? Thus, we will introduce the modal preservation operations of *generated subframe, disjoint union, p-morphic image* and *ultrafilter extension*. Moreover, the fundamental classical formation of *ultraproducts* will be used as well. All these notions will appear again and again in later sections.

Semantic structures. The structures used in the Kripke truth definition are *models* M, i.e. triples $\langle W, R, V\rangle$, where W is a nonempty set of *worlds*, R is a *binary alternative* relation on W, and V is a *valuation* assigning sets of worlds $V(p)$ to proposition letters p. The notion explicated then becomes

$$M \vDash \varphi[w] \quad : \text{`}\varphi \text{ is } \textit{true} \text{ in } M \text{ at } w\text{'}.$$

In our correspondence theory we also want to see the bare bones: a *frame* F is a couple $\langle W, R\rangle$ as above, but without a valuation. There is nothing intrinsically 'modal' about all this, of course. Frames are just the 'directed graphs' of Graph Theory.

In Sections 2.3 and 2.4, a third notion of modal structure will be required as well — intermediate, in a sense between models and frames. A *general frame* F is a couple $\langle F, \mathfrak{W}\rangle$, or alternatively, a triple $\langle W, R, \mathfrak{W}\rangle$ such that $F = \langle W, R\rangle$ is a frame, and $\mathfrak{W}$ is a set of subsets of W, closed under the formation of *complements, unions* and *modal projections*. Formally,

$$\begin{array}{ll} \text{if } X \in \mathfrak{W}, & \text{then } W - X \in \mathfrak{W} \\ \text{if } X, Y \in \mathfrak{W}, & \text{then } X \cup Y \in \mathfrak{W} \\ \text{if } X \in \mathfrak{W}, & \text{then } \pi(X) =_{\text{def}} \{w \in W \mid \exists v \in X : Rwv\} \in \mathfrak{W}. \end{array}$$

The following example illustrates the effect of restricted sets $\mathfrak{W}$. Consider the frame $\langle N, \leq\rangle$, where N is the set of natural numbers. Its modal theory contains such principles as $\Box p \to p, \Box p \to \Box\Box p$ and Geach's Axiom: together forming the logic **S4.3**. Typically left out is the McKinsey Axiom $\Box\Diamond p \to \Diamond\Box p$; as it may be falsified in some infinite alternation of $p, \neg p$: say by $V(p) = \{2n \mid n \in N\}$. But now, consider the structure $\langle N, \leq, \mathfrak{W}\rangle$, where $\mathfrak{W}$ consists of all *finite* and all *cofinite* subsets of N. It is easily checked that all three closure conditions obtain for $\mathfrak{W}$. Thus, we have a general frame here. Its logic contains the earlier one ('a fortiori'); but it also adds principles. Notably, the McKinsey Axiom can no longer be falsified, as the above 'tell-tale' valuation is no longer admissible. Thus, **S4.1** holds in this general frame, although it does not in the underlying 'full frame'. And further increases in the modal theory are possible, by restricting $\mathfrak{W}$ even more; e.g. there is even a most austere choice, viz. $\mathfrak{W} = \{\emptyset, N\}$, which yields a general frame validating the 'classical logic' with axiom $\Box p \leftrightarrow p$ — which was still invalid in the previous general frame. Thus, one single underlying frame may still generate a hierarchy of modal logics.

The original algebraic motivation for this notion (due to Thomason [1972]) will be given in Section 2.3. But here already, a direct logical reason may be given. Kripke frames are so-called 'standard models' for modal formulas, considered as second-order Π_1^1-sentences: the universal predicate quantifiers range over *all* sets of possible worlds. An intermediate possibility would have been to allow also 'general models' in the sense of Henkin [1950]: in which this second-order range may be restricted, say to some set $\mathfrak{W}$. Usually, such ranges are to be closed under certain mild conditions of definability — in order to verify reasonable forms of the universal instantiation (or 'comprehension') axiom. This, of course, is precisely what happened in the above. The uses of this notion lie partly in modal Completeness Theory, partly in modal algebra. For the moment, it will not be a major concern.

Semantic questions. Given a formal language, interpreted in certain structures, a plethora of questions arises concerning the interplay between more 'linguistic' and more 'structural' (or 'mathematical') notions. We mention only a few fundamental ones.

Arguably the 'first question' of any model theory is that concerning the relation between linguistic indistinguishability (equality of modal theories) and structural indistinguishability (isomorphism) of semantic structures. How far do the webs of language and ontology diverge? In classical logic, we know that (first-order) elementary equivalence coincides with isomorphism on the *finite* structures, but no higher up: isomorphism then becomes by far the finer sieve.

Now, the modal language on *models* behaves like the first-order language of the first translation in the introduction: nothing spectacular results. But the second-order notion seems more interesting in this respect. (Equality of second-order theories is quit'e strong: modulo the Axiom of Constructibility, it even implies isomorphism in all *countable* frames; cf. [Ajtai, 1979]). From Van Benthem [1985], which treats the analogous question for tense logic in Chapter 2.2.1, we extract

THEOREM 13. *Finite Kripke frames that are generated by a single point (cf. below) are isomorphic if and only if they possess the same modal theory. But, the countable Kripke frames* $\mathbb{Z} \odot \mathbb{Z}$ *(the integers, with each point replaced by a copy of the integers) and* $\mathbb{Q} \odot \mathbb{Z}$ *(the rationals, treated likewise) possess the same modal theory, without being isomorphic.*

In tense logic, the latter result means that the formal language cannot distinguish between locally discrete/globally discrete and locally discrete/globally dense Time. (The latter may well be that of our World.) In the context of modal logic, no such appealing interpretation is possible, whence we forego further discussion of the above result.

From now on, we will confine attention to a single theme, which again, is characteristic for much of what goes on in Model Theory.

Truth-preserving operations. In evaluating the truth of a modal formula φ at a world w we only have to consider w itself, (possibly) its R-successors, (possibly) their R-successors, etcetera. Thus, only that part of the frame is involved which is 'R-generated' by w, so to speak. In general, one never has to look beyond R-closed environments of w: an observation summed up in the following notion and result.

DEFINITION 14. M_1 ($= \langle W_1, R_1, V_1\rangle$) is a *generated submodel* of M_2 ($= \langle W_2, R_2, V_2\rangle$) (notation: $M_1 \overset{\subsetneq}{\rightarrow} M_2$) if

1. $W_1 \subseteq W_2$
2. $R_1 = R_2$ restricted to W_1,
3. $V_1(p) = V_2(p) \cap W_1$, for all proposition letters p; i.e. M_1 is an ordinary *submodel* of M_2, which has the additional feature that
4. W_1 is closed under passing to R_2-successors.

The next result is the famous 'Generation Theorem' of Segerberg [1971].

THEOREM 15. *If $M_1 \overset{\subsetneq}{\rightarrow} M_2$, then for all worlds $w \in W_1$ and all modal formulas φ, $M_1 \vDash \varphi[w]$ iff $M_2 \vDash \varphi[w]$.*

This is what happens inside a single model. When comparisons are desired between evaluation in distinct models, a more external connection is required.

DEFINITION 16. A relation C is a *zigzag connection* between two models M_1, M_2 if

1. domain $(C) = W_1$, range $(C) = W_2$,
   (a) if Cwv and $w' \in W_1$ with R_1ww', then $Cw'v'$ for some $v' \in W_2$ with R_2vv' ('forth choice')
   (b) If Cwv and $v' \in W_2$ with R_2vv', then Cww' for some $w' \in W_1$ with R_1ww' ('back choice')
2. if Cwv, then w, v verify the same proposition letters.

Starting from the basic case (3), the back-and-forth clauses ensure that evaluation of successive modalities in modal formulas yield the same results on either side:

THEOREM 17. *If M_1 is zigzag-connected to M_2 by C, then, for all worlds $w \in W_1, v \in W_2$ with Cwv, and all modal formulas φ,*

$$M_1 \vDash \varphi[w] \text{ iff } M_2 \vDash \varphi[w].$$

Notation. *$M_1 \rightleftarrows M_2$ for zigzag-connected models (by some C).*

By a result in Van Benthem [1976], the Generation Theorem and the preceding 'Zigzag Theorem' combined are characteristic for modal formulas as first-order formulas in the sense of the introduction:

THEOREM 18. *A first-order formula $\varphi(x)$ in the language with $R, P, Q, \ldots$ is logically equivalent to some modal transcription if and only if it is invariant for generated submodels and zigzag connections (in the above sense).*

For the case of pure frames, the above notions and results lead to the following three preservation results.

DEFINITION 19. F_1 is a *generated subframe* of F_2 ($F_1 \overset{\subseteq}{\to} F_2$) if

1. $W_1 \subseteq W_2$,
2. $R_1 = R_2$ restricted to W_1,
3. W_1 is R_2-closed in W_2.

In general logic, this type of situation is often described by saying that the 'converse frame' $\langle W_2, \breve{R}_2 \rangle$ is an *end extension* of $\langle W_1, \breve{R}_1 \rangle$: the added worlds all come 'at the end'.

From Theorem 15 we derive *preservation* under generated subframes:

COROLLARY 20. *If $F_1 \overset{\subseteq}{\to} F_2$, then $F_2 \vDash \varphi$ implies $F_1 \vDash \varphi$, for all modal formulas φ.*

Here '$F \vDash \varphi$' means 'φ is true in F', in the global second-order sense of the introduction: at all worlds, under all valuations.

But Theorem 15 also has an 'upward' directed moral.

DEFINITION 21. The *disjoint union* $\oplus\{F_i | i \in I\}$ of a family of frames $F_i = \langle W_i, R_i \rangle$ is the disjoint union of the domains W_i, with the obvious coordinate relations R_i.

Another direct application is *preservation under disjoint unions*:

COROLLARY 22. *If $F_i \vDash \varphi$ (all $i \in I$), then $\oplus\{F_i | i \in I\} \vDash \varphi$, for all modal formulas φ.*

Next, turning to Theorem 17, one now needs a connection between frames which can be turned into a suitable zigzag relation between models over them.

DEFINITION 23. A *zigzag morphism* from F_1 to F_2 is a function: $W_1 \to W_2$ satisfying

1. $R_1 w w'$ implies $R_2 f(w) f(w')$,
   i.e. f is an ordinary *R-homomorphism*; which has the additional backward property that
2. if $F_2 f(w) v$, then there exists $u \in W_1$ with $R_1 w u$ and $f(u) = v$.

This notion was mentioned under its current, but rather uninformative name of 'p-morphism' in the introduction. Here is one more example:

> the map from nodes to levels (counting from the top) is a zigzag morphism from the infinite binary tree (with the descendant relation) onto the natural numbers (with the usual ordering).

Notice also that injective (1-1) zigzag morphisms are even just isomorphisms.

Theorem 17 now implies the 'p-morphism' theorem of Segerberg [1971].

COROLLARY 24. *If f is a zigzag morphism from F_1 onto F_2, then, for all modal formulas $\varphi, F_1 \vDash \varphi$ implies $F_1 \vDash \varphi$.*

For more 'local' versions of these results, the reader is referred to [van Benthem, 1983].

More examples, and applications of Corollaries 20, 22, and 24 will be found in Section 2.4. A quick impression may be gained from the following sample observation (D. C. Makinson). The modal theory of any Kripke frame is either contained in the classical modal logic (characteristic axiom $\Box p \leftrightarrow p$) or the 'absurd' modal logic (characteristic axiom $\Box(p \wedge \neg p)$). For, any frame F either contains end points without R-successors, or it is *serial* ($\forall x \exists y Rxy$). In the former case, such an end point by itself forms a generated subframe, and by Corollary 20, the logic of the frame is contained in that of the subframe — which is the absurd one. In the latter case, contraction to one single reflexive point is a zigzag morphism, and by Corollary 24, the logic of the frame is contained in that of the reflexive point — which is the classical one.

We conclude by noting that these three notions are easily adapted to *general frames*, taking due precautions concerning the various sets $\mathfrak{W}_1, \mathfrak{W}_2$. Here are the three necessary additions:

In 19: add '$\mathfrak{W}_1 = \{X \cap W_1 \mid X \in \mathfrak{W}_2\}$'.

In 21: add 'the new $\mathfrak{W}_2$ remains essentially the old $\mathfrak{W}_1$' (but for the disjointness procedure used).

In 23: add the following 'continuity requirement', reminiscent of topology:

$$\text{'for all } X \in \mathfrak{W}_2, f^{-1}[X] \in \mathfrak{W}_1\text{'.}$$

These will be needed in the duality theory of Section 2.3.

Propositions and possible worlds. Another characteristic feature of modal semantics is the analogy between *propositions* and *sets of possible worlds*; as well as (moving up one stage in set-theoretic abstraction) that between *possible worlds* and *maximal sets of propositions*. Indeed, many philosophers would deny that there exist any differences here. Let us investigate.

The ideal setting here are general frames $\langle W, R, \mathfrak{W}\rangle$: the range is clearly identifiable with a collection of 'propositions' over W.

Now, if worlds are to be considered as sets of propositions, then some obvious desiderata govern the connection between a world w and propositions X, Y associated with w:

1. $X \in w$ or $Y \in w$ if and only if $X \cup Y \in w$ ('analysis')

2. $X \notin w$ if and only if $W - X \in w$ ('decisiveness').

Accordingly, one considers only subsets w of $\mathfrak{W}$ satisfying these two conditions. These are precisely the so-called *ultrafilters* on $\mathfrak{W}$.

What about the alternative relation to be imposed?

Again, a common idea is that a world v is R-accessible to w if it 'satisfies all w's modal prejudices', i.e. whenever $\Box\varphi$ is true at w, φ should be true at v. The same idea may be expressed as follows: whenever φ is true at v, $\Diamond\varphi$ should be true at w. In the present context, this becomes the following stipulation:

$$Rwv \quad \text{if for all } X \in v, \pi(X) \in w.$$

In this process, no new propositions have been created, whence the former propositions X now reappear as sets $\bar{X} = \{w \mid X \in w\}$.

These considerations motivate

DEFINITION 25. The *ultrafilter extension* $ue(G)$ of a general frame $G = \langle W, R, \mathfrak{W}\rangle$ is the general frame $\langle \mathfrak{ue}(W, \mathfrak{W}), ue(R, \mathfrak{W}), ue(\mathfrak{W})\rangle$, with

1. $\mathfrak{ue}(W, \mathfrak{W})$ is the set of all ultrafilters on $\mathfrak{W}$,

2. $\mathfrak{ue}(R, \mathfrak{W})wv$, if for each $X \in \mathfrak{W}$ such that $X \in v, \pi(X) \in w$,

3. $\mathfrak{ue}(\mathfrak{W})$ is $\{\bar{X} \mid X \in \mathfrak{W}\}$.

What this construction has done is to re-create G one level higher up in the set-theoretic air, so to speak, and some calculation will prove

THEOREM 26. *G and $ue(G)$ verify the same modal formulas.*

Still, not everything need have remained the same: the world pattern of $\langle W, R\rangle$ may differ from that of $\langle \mathfrak{ue}(W, \mathfrak{W}), \mathfrak{ue}(R, \mathfrak{W})\rangle$. First, each old world $w \in W$ generates an ultrafilter $\{X \in \mathfrak{W} \mid w \in X\}$ and, hence, a corresponding new world in $\mathfrak{ue}(W, \mathfrak{W})$. But, unless $\mathfrak{W}$ satisfies certain separation principles for worlds, different old worlds may be identified to a single new one. (In the earlier example of $\langle N, \leq, \{\emptyset, N\}\rangle$, only a single new world remains, where there used to be infinitely many!) On the other hand, the construction may also introduce worlds that were not there before. For instance, on the earlier general frame $\langle N, \leq,$ (co-)finite sets$\rangle$, the co-finite sets form an ultrafilter which induces a 'point at infinity' in the resulting ultrafilter extension. Indeed, it is easily seen that the latter consists of $\langle N, \leq\rangle$ followed by just that infinite point.

In Section 2.3, necessary and sufficient conditions will be formulated guaranteeing that a general frame is 'stable' under the construction of ultrafilter extensions. In any case, it turns out that the process stabilises after one step at the most. Now, these considerations also apply to 'full' Kripke frames.

DEFINITION 27. The *ultrafilter extension* $ue(F)$ of a frame $F = \langle W, R \rangle$ is the frame $\langle ue(W), ue(R) \rangle$, with

1. $ue(W)$ is the set of all ultrafilters on W,
2. $ue(R)wv$ if for each $X \subseteq W$ such that $X \in v, \pi(X) \in w$.

This time, Theorem 26 does not hold, however. For, it only says that the modal theory of the general frame $\langle W, R,$ power set of $W \rangle$ coincides with that of the induced general frame according to Definition 25. Now, the latter is, in general, a restriction of the full frame $\langle ue(W), ue(R) \rangle$. Hence, we can only conclude to *anti-preservation under ultrafilter extensions*:

COROLLARY 28. *If* $ue(F) \vDash \varphi$, *then* $F \vDash \varphi$, *for all modal formulas* φ.

Still, this structural notion can be made a little more familiar by connecting it with previous model-theoretic operations. First, the above-mentioned connection between old worlds and new worlds is 1-1 this time, and indeed isomorphic (consider suitable singleton sets):

THEOREM 29. F *lies isomorphically embedded in* $ue(F)$.

In general, this cannot be strengthened to 'embedded as a generated subframe'. But, another connection with the earlier preservation notions may be drawn from [van Benthem, 1979a].

THEOREM 30. $ue(F)$ *is a zigzag-morphic image of some frame* F' *which is elementarily equivalent to* F.

Proof. One expands F to $(F, X)_{X \subseteq W}$, and then passes on to a suitably saturated elementary extension, by ordinary model theory. From the latter, a canonical function from worlds to ultrafilters on F exists, which turns out to be a zigzag morphism. ■

Ultraproducts and definability. New, modally inspired notions concerning frames have been forged in the above. But old classical constructions may be considered as well. Of the various possibilities, only one is selected here, viz. the formation of ultraproducts. (For many other examples, cf. [van Benthem, 1985, Chapter I.2.1].) Its use has been indicated in the introduction already.

The basic theory (and heuristics) of the notion of 'ultraproduct' has been given in the Higher Order Logic chapter in volume 1 of this *Handbook*. (Cf. also [Chang and Keisler, 1973, Chapters 4.1 and 6.1].) We recall some of its outstanding features and uses.

DEFINITION 31. For any family of Kripke frames $\{F_i \mid i \in I\}$ with an ultrafilter U on I, the *ultraproduct* $\Pi_U F_i$ is the frame $\langle W, R \rangle$ with

1. W is the set of classes $f_\sim$, for all functions $f \in \Pi\{W_i \mid i \in I\}$, where $f_\sim$ is the equivalence class of f in the relation $f \sim g \Leftrightarrow \{i \in I \mid f(i) = g(i)\} \in U$,

2. R is the set of couples $\langle f_\sim, g_\sim \rangle$ for which $\{i \in I \mid R_i f(i) g(i)\} \in U$.

This definitional equivalence is lifted by induction to

THEOREM 32 ('Łoś Equivalence'). *For all ultraproducts, and all first-order formulas* $\varphi(x_1, \ldots, x_n)$,

$$\Pi_U F_i \vDash \varphi[f^1_\sim, \ldots, f^n_\sim] \textit{ iff } \{i \in I \mid F_i \vDash \varphi[f^1(i), \ldots, f^n(i)]\} \in U.$$

Thus, in particular, all first-order sentences φ are *preserved under ultraproducts* in the following sense:

$$\text{if } F_i \vDash \varphi (\text{all } i \in I), \quad \text{then } \Pi_u F_i \vDash \varphi.$$

Conversely, 'Keisler's Theorem' tells us that this is also enough.

THEOREM 33. *A class of Kripke frames is elementary if and only if both that class and its complement are closed under the formation of ultraproducts and isomorphic images.*

Proof. Cf. [Chang and Keisler, 1973, Chapter 6.2]. ■

A somewhat more liberal notion of definability, viz. by means of arbitrary *sets* of first-order formulas, yields so-called Δ-*elementary* classes. Here the relevant characterisation employs a special case of ultraproducts.

DEFINITION 34. An *ultrapower* $\Pi_U F$ is an ultraproduct with in each coordinate i the same frame F.

Notice that by the Łoś Equivalence, $\Pi_U F$ is *elementarily equivalent* to F, i.e. both frames possess the same first-order theory.

THEOREM 35. *A class of Kripke frames is* Δ*-elementary if and only if it is closed under the formation of ultraproducts and isomorphic images, while its complement is closed under the formation of ultrapowers.*

All these notions will be used in the modal correspondence theory of the next section. In this connection, it should be observed that, as for the other kinds of modal semantic structure, ultraproducts of *models* and of *general frames* are easily defined using the above heuristics. These will not be used in the sequel however. (Cf. [van Benthem, 1983].)

The above definability question for classical model theory leads to a clear modal task: 'to characterise the modally definable classes of Kripke frames'. In section 2.4 this matter will be investigated.

We have arrived at the interplay between classical and modal model theory, which lies at the heart of modal correspondence theory.

2.2 Correspondence I: From Modal to Classical Logic

Through the translation given in the Introduction, modal formulas may be viewed as defining constraints on the alternative relation in Kripke frames. Some of these constraints are first-order definable, others are not. Examples are presented of both, after which the former class is explored. A mathematical characterisation is given for it, in terms of ultrapowers, and methods are developed for (dis-)proving membership of the class. The limits of these methods are established as well.

First-order definability. The class of modal formulas to be studied here is defined as follows.

DEFINITION 36. **M1** consists of all modal formulas φ for which a first-order sentence α (in $R,=$) exists such that

$$F \vDash \varphi \text{ iff } F \vDash \alpha, \text{ for all Kripke frames } F.$$

Various examples of formulas in **M1** have occurred in the Introduction. For purposes of illustration, see Table 1 below.

As these are all rather easy to establish, some readers may desire a more complex example. Here it is, straight from the incompleteness Example 5 in the Introduction.

THEOREM 37. *The conjunction of the formulas $\Box p \to p, \Box\Diamond p \to \Diamond\Box p$ and $(\Diamond p \wedge \Box(p \to \Box p)) \to p$ is in* **M1**.

Proof. We shall show that this conjunction defines the same class as the classical axiom $\Box p \leftrightarrow p$, i.e. $\forall xy(Rxy \leftrightarrow x = y)$.

The argument requires several stages.

1. $\Box p \to p$ imposes reflexivity,
2. $\Diamond p \wedge \Box(p \to \Box p) \to p$ says the following:
   $\forall xy(Rxy \to \exists n \exists z_1, \ldots, z_n(Rxz_1 \wedge \ldots \wedge Rxz_n \wedge$
   $\wedge Ryz_1 \wedge \ldots \wedge Rz_nx))$.

In other words, from any R-successor y of x, one may return to x by way of some finite chain of R-successors of x. In case the chain is empty, this reduces to just: Ryx.

This (second-order!) equivalence is proved as follows (I. L. Humberstone): '$\Rightarrow$': Consider any y with Rxy. Let the *good points* be those R-successors z

Table 1.

Modal formula	Condition
$\Box p \to p$	$\forall x Rxx$
$\Box p \to \Box\Box p$	$\forall xy(Rxy \to \forall z(Ryz \to Rxz))$
$\Diamond\Box p \to \Box\Diamond p$	$\forall xy(Rxy \to \forall z(Rxz \to \exists u(Ryu \wedge Rzu)))$
$\Box(p \vee q) \to \Box p \vee \Box q$	$\forall xy(Rxy \to \forall z(Rxz \to z = y))$
$\Box(\Box p \to q) \vee \Box(\Box q \to p)$	$\forall xy(Rxy \to \forall z(Rxz \to (Ryz \vee Rzy)))$
$p \to \Box p$	$\forall xy(Rxy \to y = x)$
$\Box\bot$	$\forall x \neg\exists y Rxy$
$p \to \Box\Diamond p$	$\forall xy(Rxy \to Ryx)$

of x which can be reached from y through some finite chain (possibly empty) of R-successors of x. Then, set $V(p)$ equal to the set of all R-successors of good points. This assignment produces the following effects.

1. p is true at y (y being a successor of y, by reflexivity), and, hence, $\Diamond p$ is true at x.

2. Any R-successor of x verifying p is itself a good point, whence all *its* R-successors belong to $V(p)$.

It follows that $\Box(p \to \Box p)$ is true at x. Therefore, p itself must be true at x: i.e. x is R-successor of some good point, which was precisely to be proved.

'$\Leftarrow$': Truth of p in x is discovered by merely following the relevant chain.

3. Now, having secured *reflexivity* and '*safe return*', we can find out what the McKinsey Axiom says in the present context.

   First, notice that all R-successors of any point x may be divided up into concentric shells $S_n(x)$, where $S_n(x)$ consists of those R-successors y of x which return to x by n R- arrows (between R-successors of x) but no less. For instance, $S_0(x)$ only consists of x itself, $S_1(x)$ contains immediate R- predecessors. Notice also that, if $y \in S_{n+1}(x)$, then it must have some R-successor in $S_n(x)$.

   The McKinsey Axiom makes this whole hierarchy collapse. Set $V(p) = \cup\{S_{2n}(x) \mid n = 0, 1, 2, \ldots\}$. Then $\Box\Diamond p$ will be true at x, as follows from the above picture. For, if Rxy, and $y \in S_n(x)$, then either n is even — whence p holds at y (by definition) and so $\Diamond p$ (by reflexivity), or n is odd — whence y has an R-successor in $S_{n-1}(x)$ verifying p: which again verifies $\Diamond p$ at y.

   It follows that $\Diamond\Box p$ must be true at x. So, $\Box p$ holds at some R-successor of x. Which one? In the present situation, this can only be

x *itself.* But then again, this means that there can be no shells $S_n(x)$ with n *odd.* Thus, there is only $S_0(w) : \forall y(Rxy \to y = x)$.

4. Combining (1) and (3), the required conclusion follows: the three axioms together imply $\forall xy(Rxy \leftrightarrow y = x)$, and are obviously implied by it. ■

The very unexpectedness of this argument will have made it clear that there is a creative side to establishing correspondences.

Global and local definability. Originally, Kripke introduced frames $\langle W, R, w_0\rangle$, with a designated 'actual world' w_0. From that point of view, the study of 'local' equivalence becomes natural:

$$F \models \varphi[w] \text{ iff } \models \alpha[w],$$

where the first-order formula α has one free variable now. The reader may have noticed already that previous correspondence arguments often provide local versions as well. For instance, we had

$$\begin{array}{lll} F \models \Box p \to p[w] & \text{iff} & F \models Rxx[w] \\ F \models \Box p \to \Box\Box p[w] & \text{iff} & F \models \forall y(Rxy \to \forall z(Ryz \to Rxz))[w]. \end{array}$$

The local notion is the more informative one, in that local correspondence of φ with $\alpha(x)$ implies global correspondence of φ with $\forall x \alpha(x)$; but not conversely. Indeed, [van Benthem, 1976] contains an example of a formula in **M1** which has no local first-order equivalent at all! On the other hand, there are also circumstances under which the distinction collapses — e.g. on the *transitive* Kripke frames (W. Dziobiak; cf. [van Benthem, 1981a]).

Finally, a word of warning. Local validity of, e.g. $\Box p \to \Box\Box p$ means 'local transitivity', no more. The frame $\langle N, \{\langle 0, n\rangle \mid n \in N\} \cup \{\langle n, n+1\rangle \mid n \in N\}\rangle$ is locally transitive in 0, without being transitive.

First-order undefinability. There is a threshold of complexity below which second-order phenomena do not occur.

THEOREM 38. *All modal formulas without nestings of modal operators are in* **M1**.

Proof. Cf. [van Benthem, 1978]: a combinatorial classification suffices. ■

EXAMPLE 39. Löb's Axiom $\Box(\Box p \to p) \to \Box p$ is outside of **M1**.

Proof. It suffices to establish the following Claim: Löb's Axiom defines transitivity plus well-foundedness of the converse of the alternative relation (i.e. there are no ascending sequences $xRx_1Rx_2Rx_3, \ldots$). For, by a well-known classical compactness argument, the latter combination cannot be first-order definable (e.g. notice that it holds in $\langle N, >\rangle$, but not in its non-isomorphic ultrapowers).

First, assume that Löb's Axiom fails in F; i.e. for some V and w,

1. $\langle F, V\rangle \vDash \Box(\Box p \to p)[w]$, but

2. $\langle F, V\rangle \nvDash \Box p[w]$

Also, assume transitivity of R: we will refute the well-foundedness of $\breve{R}$, by constructing an endless ascending sequence of worlds $wRw_1Rw_2\ldots$.

Step 1: Chose any w_1 with Rww_1 where p fails (by (2)). By (1), $\Box p \to p$ is true at w_1, whence $\Box p$ fails again.

Step 2: chose any w_2 with Rw_1w_2 where p fails. By (1) and transitivity, $\Box p \to p$ is true at w_2, etcetera: an endless sequence is on its way.

Next, failure of either of the two relational conditions results in failure of Löb's Axiom. If transitivity fails, say $Rwv, Rvu, \neg Rwu$, then $V(p) = W - \{v, u\}$ verifies $\Box(\Box p \to p)$ at w, while falsifying $\Box p$.

If well-foundedness fails, say $wRw_1Rw_2, \ldots$, then $V(p) = W - \{w, w_1, w_2, \ldots\}$ produces the same effect. ■

More complex undefinability arguments will be discussed later on.

First-order definability and ultraproducts. Modal formulas could be regarded as Π^1_1-sentences, witness the Introduction. Now, for the latter sentences, ultraproducts provide the touchstone for first-order definability:

THEOREM 40. *A Π^1_1-sentence in $R, =$ is first-order definable if and only if it is preserved under ultraproducts.*

Proof. '$\Rightarrow$': This follows from the Łoś Equivalence (cf. Section 2.1).

'$\Leftarrow$': Consider a typical such sentence:

$$\forall P_1 \ldots \forall P_n \varphi(P_1, \ldots, P_n, R, =) \quad (\varphi \text{ first-order}).$$

Clearly it is preserved under *isomorphisms* (and so is its negation). Moreover, its negation (a 'Σ^1_1-sentence') is preserved under *ultraproducts* (cf. [Chang and Keisler, 1973, Chapter 4.1], for the easy argument). So, given the assumption on the sentence itself, Keisler's Theorem (33) applies. ■

COROLLARY 41. *A modal formula is in* **M1** *if and only if it is preserved under ultraproducts.*

A second application says that no generalisation of our topic is obtained by allowing arbitrary *sets* of defining first-order conditions.

COROLLARY 42. *If a modal formula has a Δ-elementary definition, it has an elementary definition.*

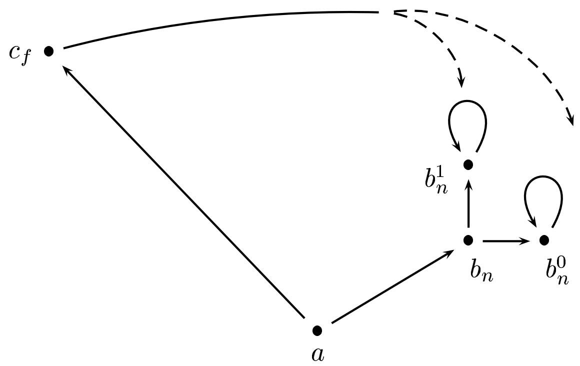
Proof. Δ-elementary classes are closed under the formation of ultraproducts. ■

This characterisation of **M1** is rather aspecific, as it holds for *all* Π^1_1-sentences. Later on, we will exploit the specifically *modal* character of our formulas to do better. Moreover, the characterisation is rather abstract, as ultraproducts are hard to visualise. Therefore, we now turn to more concrete methods for separating formulas inside **M1** from those outside.

Formulas beyond **M1***: Compactness and Löwenheim–Skolem arguments.* In practice, nonfirst-order definability often shows up in failure of the Compactness and Löwenheim–Skolem theorems. The first was involved in the example of Löb's Axiom, the second will be presented now.

EXAMPLE 43 (McKinsey's Axiom). $\Box\Diamond p \to \Diamond\Box p$ is outside of **M1**.

Proof. Consider the following uncountably infinite Kripke frame $F = \langle W, R\rangle$:



$$
\begin{array}{rcl}
W & = & \{a\} \cup \{b_n, b_n^0, b_n^1 \mid n \in N\} \cup \{c_f \mid f : N \to \{0,1\}\} \\
R & = & \{\langle a, b_n\rangle, \langle b_n, b_n^0\rangle, \langle b_n, b_n^1\rangle, \langle b_n^0, b_n^0\rangle, \langle b_n^1, b_n^1\rangle \mid n \in N\} \cup \\
 & & \{\langle a, c_f\rangle \mid f : N \to \{0,1\}\} \cup \{\langle c_f, b_n^{f(n)}\rangle \mid n \in N, f : N \to \{0,1\}\}.
\end{array}
$$

We observe two things.

1. $F \models \Box\Diamond p \to \Diamond\Box p$.

Thanks to the presence of the reflexive endpoints b_n^0, b_n^1, the validity of the McKinsey Axiom is obvious everywhere, except for a.

So, suppose that, under some valuation $V, \Box\Diamond p$ is true at a. By assumption, $\Diamond p$ is true at each b_n, and hence p is true at b_n^0 or b_n^1. Now, pick any function $f : N \to \{0,1\}$ such that $b_n^{f(n)}$ is a p-world (each $n \in N$). Then $\Box p$ holds at c_f, and hence $\Diamond\Box p$ at a.

By the downward Löwenheim–Skolem theorem, F possesses a *countable* elementary substructure F' whose domain contains (at least) a, b_n, b_n^0, b_n^1 (all $n \in N$). As F is *uncountable*, many worlds (c_f) must be missing in

W'. Fix any one of these, say c_{f_0}. Notice, for a start, that c_{1-f_0} cannot be in W' either. (For, the existence of 'complementary' c-worlds is first-order expressible; and F' verifies the same first-order formulas at each of its worlds as F.) Now, setting

$$V(p) = \{b_n^{f_0(n)} \mid n \in N\}$$

will verify $\Box\Diamond p$ at a, while falsifying $\Diamond\Box p$. Thus, we have shown

2. $F' \not\models \Box\Diamond p \to \Diamond\Box p$.

We may conclude that the McKinsey Axiom is not first-order definable — not being preserved under elementary subframes. ■

In *practice*, failure of Löwenheim–Skolem or compactness properties is an infallible mark of being outside of **M1**. The reader may also think this to be the case in *theory*, by the famous Lindström Theorem. (Cf. Volume 1, chapters by Hodges or van Benthem and Doets.) But there is a little-realised problem: the Lindström Theorem does not work for languages with a fixed finite vocabulary (cf. [van Benthem, 1976]). In our case of $R, =$, there do exist proper extensions of predicate logic satisfying both the Löwenheim and compactness properties. These are not *modal* examples, however — and it may well be the case, for all we know, that a modal formula φ belongs to **M1** if and only if the logic obtained by adding φ to the first-order predicate logic in $R, =$ as a propositional constant has the Löwenheim and compactness properties. Indeed, up till now, all undefinability arguments (including the above) have always been found reducible to *compactness* arguments alone.

The final characterisation of **M1**. Corollary 41 may be improved by noting the following fact about Kripke frames, connecting the modal and classical notions of Section 2.1.

LEMMA 44. $\Pi_U F_i \overset{\subsetneq}{\to} \Pi_U \oplus \{F_i \mid i \in I\}$.

Thus, ultraproducts are generated subframes of suitable ultrapowers.

A second idea comes from the preceding section: outside of **M1**, we encountered non preservation under *elementary equivalence*, a notion tied up with ultrapowers by the Keisler–Shelah Theorem (cf. [Chang and Keisler, 1973, Chapter 6.1]). We arrive at the main result of [van Benthem, 1976].

THEOREM 45. *(i) A modal formula is in* **M1** *if and only if (ii) it is preserved under ultrapowers if and only if (iii) it is preserved under elementary equivalence.*

Proof. (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii) are immediate. (ii) $\Rightarrow$ (i): If φ is preserved under ultrapowers, then, by Lemma 44, it is also preserved under ultraproducts — because disjoint unions preserve modal truth (Corollary 22). Now apply Corollary 41. ■

Again, this insight saves us some spurious generalisations. Besides 'Δ-elementary', there are two more levels in the definability hierarchy

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

A Σ-elementary class is defined by an infinite *disjunction* of first-order sentences (Δ-elementary classes by infinite *conjunctions*). The prime example of this phenomenon is *finiteness.* $\Sigma\Delta$-elementary classes arise from infinite disjunctions of infinite conjunctions, or vice versa: both cases (and all purported 'higher' ones) collapse — and the hierarchy stops here, even in classical logic. The reason lies in the simple observation that a class of frames is $\Sigma\Delta$-elementary if and only if it is closed under *elementary equivalence.*

But the preceding result has a

COROLLARY 46. *Modal formulas are either elementary, or essentially higher-order.*

Unfortunately, even this better characterisation does not yield much *effective* information concerning the members of **M1**. For, there are no syntactic criteria for preservation under ultrapowers. From [van Benthem, 1983], we will cite the catalogue of what little we know.

DIGRESSION 47.

1. Π^1_1-sentences in $R, =$ of the purely universal form

   $$\forall P_1 \ldots \forall P_m \forall x_1 \ldots \forall x_n \varphi \qquad (\varphi \text{ quantifier-free})$$

   are preserved under ultraproducts. This tells us that $p \to \Box p$, i.e.

   $$\forall P \forall x (Px \to \forall y (Rxy \to Py))$$

   must be in **M1**: but that was clear without such heavy artillery.

2. Π^1_1-sentences in $R, =$ of the universal-existential form

   $$\forall P_1 \ldots \forall P_m \exists x_1 \ldots \exists x_n \varphi \qquad (\varphi \text{ quantifier-free})$$

   are preserved under ultrapowers. This is of no help whatsoever, as modal formulas have at least one *universal* first-order quantifier ($\forall x$).

3. Further presents will not be forthcoming: *any* Π^1_1-sentence in $R, =$ is logically equivalent to one of the form

   $$\forall P_1 \ldots \forall P_m \forall x_1 \ldots \forall x_n \exists y_1 \ldots \exists y_n \varphi \qquad (\varphi \text{ quantifier-free})$$

   So, all complexity occurs at this level already.

Thus, other ways are to be developed for describing **M1** effectively.

The method of substitutions. There is a common syntactic pattern to many examples of first- order definable modal formulas: certain antecedents, in combination with certain consequents enable one to 'read off' equivalents. Starting from the earlier examples $\Box p \to p, \Diamond\Box p \to \Box\Diamond p$, one may notice successively that conjunctions and disjunctions are admissible as well; as long as one avoids $\Box\Diamond$ or $\Box(\ldots \vee \ldots)$ combinations to the left.

A typical instance is the following result from [Sahlqvist, 1975]:

THEOREM 48. *Modal formulas $\varphi \to \psi$ are in* **M1***, provided that*

1. *φ is constructed from the forms $p, \Box p, \Box\Box p, \ldots, \bot, \top$, using only $\wedge, \vee$ and $\Diamond$, while*
2. *φ is constructed from proposition letters, $\bot, \top$, using $\wedge, \vee, \Diamond$ and $\Box$.*

This theorem accounts for cases such as

$$\Diamond(p \wedge \Box q) \to \Box(p \vee \Diamond p \vee q)$$

which defines

$$\forall xy(Rxy \to \forall z(Rxz \to (z = y \vee Rzy \vee Ryz))).$$

Proof. The heuristics of the Introduction works: for each 'minimal verification' of the antecedent, the consequent must hold. For further technical information (e.g. the *monotonicity* of the consequent is vital too), cf. [van Benthem, 1976], which also contains generalisations of the theorem. ■

That $\Box\Diamond$ is fatal, is shown by the McKinsey Axiom. The Fine Axiom $\Diamond\Box(p \vee q) \to \Diamond(\Box p \vee \Box q)$ does the same for $\Box(\ldots \vee \ldots)$. Finally, the Löb Axiom (in the equivalent form $\Diamond p \to \Diamond(p \wedge \Box\neg p)$) demonstrates the danger of 'negative' parts in the consequent. Thus, in a sense, we have a 'best result' here.

Notice that the class described is rather typical for modal axioms, which often assume this implicational form. Indeed, the most characteristic modal axioms are even simply *reduction principles* of the form

$$(\text{modal operators})\ p \to (\text{modal operators})\ p.$$

THEOREM 49. *A modal reduction principle is in* **M1** *if and only if it is of one of the following four types:*

1. *$\vec{M}p \to \Box \ldots \Box\Diamond \ldots \Diamond p$,*
2. *$\Diamond \ldots \Diamond\Box \ldots \Box p \to \vec{M}p$,*
3. *$\Box \ldots$ (i times) $\ldots \Box\vec{M}p \to \vec{N}\vec{M}p$ (where length $(\vec{N}) = i$),*

4. $\vec{N}\vec{M}p \to \Diamond \ldots$ *(i times)* $\ldots \Diamond\vec{M}p$ *(where length* $(\vec{N}) = i$*).*

Proof. Cf. [van Benthem, 1976] for the rather laborious argument. ■

Thus at least, important parts of **M1** have been classified. This particular theorem finishes a project begun in [Fitch, 1973].

A general method of proof for Theorem 48 consists of the method of substitutions, introduced in the introduction. Here we shall merely illustrate how it works: a justification may be found in [van Benthem, 1983].

EXAMPLE 50. Write $\Diamond\Box p \to \Box\Diamond p$ as

$$\forall P \forall x(\exists y(Rxy \wedge \forall z(Ryz \to Pz)) \to \forall u(Rxu \to \exists v(Ruv \wedge Pv))).$$

Rewrite this to the equivalent

$$\forall xy(Rxy \to \forall P(\forall z(Ryz \to Pz) \to \forall u(Rxu \to \exists v(Ruv \wedge Pv)))).$$

Substitute for $P : \lambda z.Ryz$, to obtain

$$\forall xy(Rxy \to (\forall z(Ryz \to Ryz) \to \forall u(Rxu \to \exists v(Ruv \wedge Ryv)))).$$

This is equivalent to

$$\forall xy(Rxy \to \forall u(Rxu \to \exists v(Ruv \wedge Ryv))),$$

i.e. *directedness (confluence).*

Write $\Diamond(p \wedge \Box q) \to \Box(p \vee \Diamond p \vee q)$ as

$$\forall xy(Rxy \to \forall P((Py \wedge \forall z(Ryz \to Qz)) \to \forall u(Rxu \to (Pu \vee \\ \vee \exists v(Ruv \wedge Pv) \vee Qu)))).$$

Substitute for $P : \lambda z{\cdot}y = z$, and for $Q : \lambda z.Ryz$, to obtain (an equivalent of) the earlier *connectedness.*

Write $\Diamond(p \wedge \Box p) \to p$ as

$$\forall xy(Rxy \to \forall P((Py \wedge \forall z(Ryz \to Pz)) \to Px)).$$

Substitute for $P : \lambda z{\cdot}y = z \vee Ryz$, to obtain (an equivalent of)

$$\forall xy(Rxy \to (Ryx \vee y = x)).$$

Write $\Box\Box p \to \Box p$ as

$$\forall x \forall P(\forall y(Rxy \to \forall z(Ryz \to Pz)) \to \forall u(Rxu \to Pu)).$$

Substitute for $P : \lambda z \cdot R^2xz$; i.e. $\lambda z \cdot \exists v(Rxv \wedge Rvz)$, to obtain (modulo logical equivalence)

$$\forall x \forall u(Rxu \to \exists v(Rxv \wedge Rvu)),$$

i.e., *density* of the alternative relation.

In general, substitutions will be disjunctions of forms $R^n yz (n = 0, 1, 2, \ldots)$; the cases 0, 1 standing for $=, R$, respectively.

Despite these advances, the range of the method of substitutions has it limits. To see this, here is an example of a formula in **M1** with a quite different spirit.

EXAMPLE 51. The conjunction of the **K4.1** axioms, i.e. $\Box p \to \Box\Box p$, $\Box\Diamond p \to \Diamond\Box p$ is in **M1**.

Proof. $\Box p \to \Box\Box p$ defined transitivity and, therefore, it suffices to prove the following

CLAIM. On the *transitive* Kripke frames, McKinsey's Axiom defines *atomicity*:

$$\forall x \exists y(Rxy \wedge \forall z(Ryz \to z = y)).$$

From right to left, the implication is clear. From left to right, however, the argument runs deeper.

Assume that F is a transitive frame, containing a world $w \in W$ such that

$$\forall y(Rwy \to \exists z(Ryz \wedge z \neq y)).$$

Using some suitable form of the Axiom of Choice (it is as serious as this ...), find a subset X of w's R-successors such that

1. $\forall y \in W(Rwy \to \exists z \in X Ryz)$
2. $\forall y \in W(Rwy \to \exists z \in (W - X)Ryz)$.

Setting $V(p) = X$ then falsifies the McKinsey Axiom at w. ■

This complexity is unavoidable. We can, for example, prove

THEOREM 52. $(\Box p \to \Box\Box p) \wedge (\Box\Diamond p \to \Diamond\Box p)$ *is not equivalent to any conjunction of its first-order substitution instances.*

Proof. Here is where the earlier general frame $\langle N, \leq,$ finite and cofinite sets$\rangle$ comes in. First, an ordinary model-theoretic

OBSERVATION. The finite and cofinite sets of natural numbers are precisely those first-order definable in $\langle N, \leq\rangle$, possibly using parameters.

Now, it was noticed already in Section 2.1 that the above formula holds in this general frame — and hence so do all its first-order substitution instances. But the latter also hold in the full frame $\langle N, \leq \rangle$. So, if our formula were defined by them, it would also hold in the full frame: which it does not. ■

So, although he method of substitutions carves out a large, and important part of **M1**, it does not fully describe the latter class.

The complexity of **M1**. The method of substitutions describes a part of **M1** which may even be shown to be *recursively enumerable* (cf. [van Benthem, 1983]). But **M1** overflowed its boundaries. Indeed, there are reasons to believe that **M1** is not recursively enumerable — probably not even arithmetically definable. For, in the general case of Π_1^1-sentences, we know

THEOREM 53. *First-order definability of* Π_1^1*-sentences is not an arithmetical notion.*

Proof. (Cf. [van Benthem, 1983] or the Higher Order Logic Chapter in Volume 1 of this *Handbook*.) ■

Other topics. Various other questions had to be omitted here. At least, one example should be mentioned, viz. that of *relative correspondences.* On several occasions, a restriction to *transitive* Kripke frames produced interesting shifts: global and local first-order definability collapse, the McKinsey Axiom becomes elementary, etc. A sample result is in [van Benthem, 1976].

THEOREM 54. *On the transitive Kripke frames, all modal reduction principles are first-order definable.*

Thus, 'pre-conditions' on the alternative relation are worth considering. In areas such as tense logic, our temporal intuitions even *require* them.

2.3 Modal Algebra

An alternative to Kripke semantic structures is offered by so-called '*modal algebras*', in which the modal language may be interpreted as well. The realm of modal algebras has its own mathematical structure, with *subalgebras, direct products* and *homomorphic images* as key notions. Now, back-and-forth connections may be established between these two realms, through the *Stone Representation.* A categorial parallel emerges between the above triad of notions and the basic triad of Section 2.1: zigzag-morphic images, disjoint unions and generated subframes, respectively. Moreover, the earlier 'possible worlds construction' for ultrafilter extensions will be seen to arise naturally from the Stone Representation.

The algebraic perspective. As in other areas of logic, the modal propositional language may also be interpreted in *algebraic* structures. These assume the form of a Boolean Algebra (needed to interpret the propositional base) enriched with a unary operation, in order to capture the modal operator.

DEFINITION 55. A *modal algebra* is a tuple

$$\mathfrak{A} = \langle A, 0, 1, +, ', * \rangle,$$

where $\langle A, 0, 1, +, ' \rangle$ is a Boolean Algebra and $*$ is a unary operator satisfying the equations

1. $(x + y)^* = x^* + y^*$
2. $0^* = 0$.

Notice that $*$ corresponds to possiblity ($\Diamond$): the necessity choice would have yielded equations

$1'$. $(x \cdot y)^* = x^* \cdot y^*$

$2'$. $1^* = 1$.

This algebraic perspective at once yields a completeness result.

THEOREM 56. *A modal formula is derivable in the minimal modal logic* **K** *if and only if it receives value 1 in all modal algebras under all assignments.*

The concept of evaluation at the back of this goes as follows. Let V assign A-values to proposition letters. Then, V may be lifted to all formulas through the recursive clauses

$$\begin{aligned} &V(\neg\varphi) = V(\varphi)' \\ &V(\varphi \vee \psi) = V(\varphi) + V(\psi) \\ &V(\Diamond\varphi) = V(\varphi)^*, \text{ etc.} \end{aligned}$$

Thus, a modal formula is read as a 'polynomial' in $', +, *$.

The *proof* of the completeness Theorem 56 comes cheap. First, one shows by induction on the length of proofs that all **K**-theorems are 'polynomials identical to 1'. Conversely, one considers the so-called *Lindenbaum Algebra* of the modal language, whose elements are equivalence classes of **K**-provably equivalent modal formulas, with operations defined in the obvious way through the connectives. The value 1 in this algebra is awarded to all and only the **K**-theorems: hence non- theorems are disqualified as polynomials identical to 1.

Such uses of modal algebra are a joy to some (cf. [Rasiowa and Sikorski, 1970]); to others they show that the algebraic approach is merely 'syntax in disguise'. After all, the above result may be viewed as a re-axiomatisation of **K**, no more. For instance, notice that the hard work in the usual (Henkin type) model-theoretic completeness theorems consists in showing that non-theorems can be refuted in *set-theoretic* (Kripke)-models. To put this into a slogan, which will become fully comprehensible at the end of this chapter:

HENKIN = LINDENBAUM + STONE.

Nevertheless, the algebraic perspective has further uses, which are being discovered only gradually. First, notice that it offers a more general framework than Kripke semantics. For the above Lindenbaum construction to work, one only needs the principle of Replacement of Equivalents; i.e. modally, closure under the rule

$$\text{if } \vdash \varphi \leftrightarrow \psi, \text{ then } \vdash \Diamond\varphi \leftrightarrow \Diamond\psi.$$

(Algebraically, this just amounts to an identity axiom.)

The above additional equations represent optional further choices.

But even in the realm of the above modal algebra, there exists a whole discipline of universal algebraic notions and results, which turn out to be applicable to modal logic in surprising ways. Two instructive references are [Goldblatt, 1979] and [Blok, 1976]. Here we shall only skim the surface, taking what is needed for the modal definability results of Section 2.4. Thus, we shall need the following three fundamental algebraic notions.

DEFINITION 57. $\mathfrak{A}_1$ is a modal *subalgebra* of $\mathfrak{A}_2$ if $A_1 \subseteq A_2$, and the operations of $\mathfrak{A}_2$ coincide with those of $\mathfrak{A}_1$ on A_1.

DEFINITION 58. The *direct product* $\Pi\{\mathfrak{A}_i \mid i \in I\}$ of a family of modal algebras $\{\mathfrak{A}_i \mid i \in I\}$ consists of all functions in the Cartesian product $\Pi\{A_i \mid i \in I\}$, with operations defined component-wise:

$$f + g = (f(i) +_i g(i))_i, \quad f^* = (f(i)^*_i)_i; \text{ etc.}$$

DEFINITION 59. A function f is a *homomorphism* from $\mathfrak{A}_1$ to $\mathfrak{A}_2$ if it respects all operations:

$$f(a +_1 b) = f(a) +_2 f(b), \quad f(a^{*_1}) = f(a)^{*_2}; \text{ etc.}$$

These three operations are fundamental in algebra because they characterise algebraic *equational definability.* This is the content of 'Birkhoff's Theorem':

A class of algebras is defined by the validity of a certain set of algebraic equations (under all assignments) if and only if that class is closed under the formation of subalgebras, direct products and homomorphic images. (For a proof, cf. [Grätzer, 1968].) There is much more to Universal Algebra, of course, but this is what we shall need in the sequel.

Kripke frames induce modal algebras. In order to tap the above resources, a systematic connection is needed between the earlier semantic structures and modal algebras.

To begin with, each Kripke frame $F = \langle W, R\rangle$ gives rise to the following modal algebra

$$A(F) = \langle P(W), \varnothing, W, \cup, -, \pi\rangle$$

where π is the *modal projection* of 2.1:

$$\pi(X) = \{w \in W \mid \exists v \in X Rwv\} \quad (X \subseteq W).$$

As for truth of modal formulas, it is immediate that a modal formula φ is true in F if and only if its corresponding modal equation $a(\varphi)$ is identical to 1 in the algebra $A(F)$. For instance, truth of

$$\Diamond\Box(p \vee q) \to \Diamond(\Box p \vee \Box q),$$

or equivalently

$$\neg\Diamond\neg\Diamond\neg(p \vee q) \vee \Diamond(\neg\Diamond\neg p \vee \neg\Diamond\neg q)$$

is equivalent to the validity of the identity

$$(x+y)^{\prime *\prime *\prime} + (x^{\prime *\prime} + y^{\prime *\prime})^* = 1.$$

Thus, A maps single Kripke frames to modal algebras. But what happens to the characteristic modal connections between frames, as in Section 2.1? We shall take them one by one.

First, if F_1 is a *generated subframe* of F_2, then the obvious restriction map sending $X \subseteq W_2$ to $X \cap W_1$ is a modal *homomorphism* from $A(F_2)$ *onto* $A(F_1)$. (The key observation is that R_2-closure of W_1 guarantees homomorphic respect for the projection operator π.) Next, the algebra induced by a *disjoint union* $\oplus\{F_i \mid i \in I\}$ is isomorphic, in a natural way, to the *direct product* $\Pi\{A(F_i) \mid i \in I\}$. One simply associates a set X of worlds in the former with the function $(X \cap W_i)_{i \in I}$.

Finally, and this happy ending will be predictable by now, if F_2 is a zigzag-morphic image of F_1 through f, then the stipulation

$$A(f)(X) =_{\text{def}} f^{-1}[X]$$

defines an isomorphism between $A(F_2)$ and a *subalgebra* of $A(F_1)$. (This time, the two relational clauses in the definition of 'zigzag morphism' ensure that $A(f)$ respects projections.) Notice the reversal in direction in the latter case: this is a common phenomenon in these 'categorial connections'.

Modal algebras induce Kripke structures. There is a road back. Conversely, modal algebras may be 'represented' as if they had come from an underlying base frame. The idea of this so-called *Stone Representation* is as follows. (It is due to Jónsson and Tarski around 1950.)

Worlds w are to be created such that an element a in the algebra may be thought of as the set of w 'in a'. But then, the desired correspondence between algebraic and set-theoretic operations becomes:

no set w is in 0, all sets w are in 1,
w is in $a+b$ iff w is in a or w is in b,
w is in a' iff w is not in a.

Thus, as w searches through A 'where it belongs', it picks out a set X such that

$$\begin{array}{llll} 0 \notin X, & 1 \in X, & & \\ a+b \in X & \text{iff} & a \in X & \text{or} \quad b \in X, \\ a' \in X & \text{iff} & a \notin X. & \end{array}$$

Such sets X are called *ultrafilters* on $\mathfrak{A}$. Thus, let

$$W(\mathfrak{A}) = \text{ all ultrafilters on } \mathfrak{A}.$$

A suitable alternative relation may be found through the same motivation as in Section 2.1.

$$\langle w, v\rangle \in R(\mathfrak{A}) \quad \text{iff} \qquad \text{for each } a \in A, \text{ if } a \in v, \text{ then } a^* \in w.$$

So, each modal algebra $\mathfrak{A}$ induces a Kripke frame

$$F(\mathfrak{A}) = \langle W(\mathfrak{A}), R(\mathfrak{A})\rangle.$$

This time, truth in $\mathfrak{A}$ and truth in $F(\mathfrak{A})$ need not correspond, however. For, $F(\mathfrak{A})$ may harbour many more sets of worlds than just those corresponding to the elements a of the algebra — and hence it contains additional potential falsifiers. Thus, the implication goes only one way. The equation $t_1 = t_2$ is valid in $\mathfrak{A}$, where the polynomials t_1, t_2 correspond to the modal formulas φ_1, φ_2, when $\varphi_1 \leftrightarrow \varphi_2$ is true in $F(\mathfrak{A})$. A complete equivalence is only restored by changing $F(\mathfrak{A})$ to the *general frame*

$$F(\mathfrak{A}) = \langle W(\mathfrak{A}), R(\mathfrak{A}), \mathfrak{W}(\mathfrak{A})\rangle,$$

where $\mathfrak{W}(\mathfrak{A})$ consists of all sets of the form

$$\{w \in W(\mathfrak{A}) \mid a \in w\} \quad (a \in A).$$

So, what we now get is a two-way connection between *modal algebras* and *general frames* — and here lies the genesis of the latter notion. Two ways; for, it is easily seen that all previous insights about the mapping A apply equally well to general frames, instead of merely 'full' frames.

Again, the interest of the present connection may be gauged by seeing what happens to the three fundamental algebraic operations when translated through F into Kripke-semantic terms.

First, if $\mathfrak{A}_1$ is a modal *subalgebra* of $\mathfrak{A}_2$, then the obvious restriction map sending ultrafilters w on $\mathfrak{A}_2$ to ultrafilters $w \cap A_1$ on $\mathfrak{A}_1$ is a *zigzag morphism* from $F(\mathfrak{A}_2)$ *onto* $F(\mathfrak{A}_1)$.

Next, the *direct product* of a family $\{\mathfrak{A}_i \mid i \in I\}$ has an F-image containing the *disjoint union* $\oplus\{F(\mathfrak{A}_i) \mid i \in I\}$. No isomorphism need obtain, however: a slight flaw in our correspondence.

But finally, if $\mathfrak{A}_2$ is a *homomorphic image* of $\mathfrak{A}_1$ through f, then the map $F(f)$, defined by setting

$$F(f)(w) =_{\text{def}} f^{-1}[w],$$

sends $\mathfrak{A}_2$-ultrafilters to $\mathfrak{A}_1$-ultrafilters, in such a way that it embeds $F(\mathfrak{A}_2)$ isomorphically as a *generated subframe* of $F(\mathfrak{A}_1)$.

Back and forth. So far, so good. Modal algebras induce general frames, and these, in their turn, induce modal algebras. But, what happens on a return-trip?

One case is simple, by construction:

THEOREM 60. $A(F(\mathfrak{A}))$ *is isomorphic to* $\mathfrak{A}$.

The converse direction is more difficult. $(F(A(G))$ need not be isomorphic to F, for general frames G. This is precisely what we noted in connection with 'possible world constructions' in Section 2.1. But, as was announced there, it can be ascertained which conditions on general frames G do guarantee such an isomorphism.

DEFINITION 61. A general frame $G = \langle W, R, \mathfrak{W} \rangle$ is *descriptive* if it satisfies *Leibniz' Principle* for *identity*:

1. $\forall xy \in W(x = y \leftrightarrow \forall Z \in \mathfrak{W}(x \in Z \leftrightarrow y \in Z))$

as well as *Leibniz' Principle* for *alternatives*:

2. $\forall xy \in W(Rxy \leftrightarrow \forall Z \in \mathfrak{W}(y \in Z \rightarrow x \in \pi(Z)))$.

Moreover, it should satisfy *Saturation*:

3. each subset of $\mathfrak{W}$ with the finite intersection property has a non-empty total intersection.

The following basic result is in [Goldblatt, 1979].

THEOREM 62. $F(A(G))$ *is isomorphic to* G *if and only if* G *is descriptive.*

The standard examples of descriptive frames are the general frames derived from *Henkin models* in modal completeness proofs, by taking for $\mathfrak{W}$ the range of modally definable sets of worlds. It may also be noticed that general frames G which are themselves of the form $F(\mathfrak{A})$ are always descriptive. Thus, for certain theoretical purposes, the 'proper' bijective correspondence may be said to be that between modal algebras and descriptive frames, which are 'stable' under the possible worlds construction described in Section 2.1.

The categorial connection. The above connections between modal algebras and Kripke structures run deeper than might appear at first sight. The general picture is that of two mathematical worlds, or 'categories', which turn out to be quite similar in structure:

$$\langle \text{Modal algebras, homomorphisms into} \rangle$$
$$\langle \text{General frames, zigzag morphisms into} \rangle.$$

The earlier considerations may be summed up in the following two schemata:

$$\begin{tikzcd} G_1 \arrow[r, "f"] \arrow[d, dash] & G_2 \arrow[d, dash] \\ A(G_1) & A(G_2) \arrow[l, "A(f)"'] \end{tikzcd} \qquad \begin{tikzcd} \mathfrak{A}_1 \arrow[r, "f"] \arrow[d, dash] & \mathfrak{A}_2 \arrow[d, dash] \\ F(\mathfrak{A}_1) & F(\mathfrak{A}_2) \arrow[l, "F(f)"'] \end{tikzcd}$$

So, A, F are what a category theorist would call 'contravariant' functors. Therefore, information concerning the one category may sometimes be transferred to the other. Thus, a 'categorial transfer' arises, of which we mention a few phenomena. (The following passage can be skipped by readers unfamiliar with Category Theory or Universal Algebra).

The category of modal algebras has among its internal limit constructions the formation of *terminals* (viz. the degenerate single point algebras) and *pull-backs*. Hence, it is closed under *finite limits* in general. Through A, F, we may derive that the category of general frames is closed under *finite co-limits*, specifically under *initials* (allowing the *empty* frame) and *push-outs*. (In this connection, the 'adjointness' behaviour of A, F may be investigated.) The preservation behaviour of modal formulas under such limit constructions remains to be studied.

An algebraically well-motivated notion is that of a *free algebra*. What corresponds to these in the realm of general frames? A surprising connection with modal completeness theory appears. The Stone representations of free algebras are essentially *Henkin general frames* (proposition letters correspond to free generators of the algebra). The latter structures were characterised semantically in [Fine, 1975], in terms of certain 'universal embedding' properties with respect to zigzag morphisms. This turns out to follow directly, as the dual of the 'homomorphic extension' definition of free algebras.

Our final example concerns another algebraic classic, the notion of a *subdirectly irreducible* modal algebra (used with great versatility in [Blok, 1976]). These turn out to correspond almost (not quite) to rooted general frames whose domain consists of one root world together with its R-successors, their R-successors, etcetera. The famous Birkhoff Theorem stating that

> Every (modal) algebra is a subdirect product of subdirectly irreducibles,

may then be compared with the simple Kripke-semantic observation that

> Every general frame is a zigzag-morphic image of the disjoint union of its rooted generated subframes.

These examples will have made it clear how the categorial connection between modal algebra and possible worlds semantics can be a very rewarding perspective.

2.4 From Classical to Modal Logic

Reversing the direction of the earlier correspondence study (Section 2.2), there arises

DEFINITION 63. **P1** is the set of all first-order sentences in $R, =$ for which a modal formula exists defining the same class of Kripke frames.

All earlier examples of formulas in **M1** also provide examples for **P1**, of course. Therefore, here are some more general results straightaway.

Some methods exist for *proving* the existence of modal definitions.

THEOREM 64. *Each first-order sentence of the form $\forall x U \varphi$, where U is a (possibly empty) sequence of restricted universal quantifiers, of the form*

$$\forall u(Rvu \to \quad \textit{(with } u, v \textit{ distinct)}$$

followed by a matrix φ of atomic formulas $u = v, Ruv$ combined through $\wedge, \vee$, belongs to **P1**.

Proof. The relevant combinatorial argument is based on the heuristics explained in the introduction. Cf. [van Benthem, 1976]. ■

Some examples of formulas of this type are

$$\text{reflexivity: } \forall x Rxx, \text{ transitivity: } \forall x \forall y (Rxy \to \forall z (Ryz \to Rxz))$$

and

$$\text{connectedness: } \forall x \forall y (Rxy \to \forall z (Rxz \to (Rzy \vee Ryz))).$$

Disproving definability proceeds through counter-examples to preservation behaviour.

EXAMPLE 65.

1. $\exists x Rxx$ is outside of **P1**.

It holds in $\langle \{0, 1\}, \{\langle 1, 1\rangle\}\rangle$; but not in its generated subframe $\langle \{0\}, \varnothing\rangle$.

2. $\forall x \forall y Rxy$ is outside of **P1**.

It is preserved under generated subframes, but not under disjoint unions. On $\langle\{0\},\{\langle 0,0\rangle\}\rangle$ and $\langle\{1\},\{\langle 1,1\rangle\}\rangle$, the relation is universal; but not on $\langle\{0,1\},\{\langle 0,0\rangle,\langle 1,1\rangle\}\rangle$.

3. $\forall x \neg Rxx$ is outside of **P1**.

It is preserved under generated subframes and disjoint unions; but not under zigzag-morphic images, witness the Introduction.

4. $\forall x \exists y (Rxy \wedge Ryy)$ is outside of **P1**.

It is preserved under all three operations mentioned up till now, but not inversely under the formation of ultrafilter extensions. It can be shown to hold in $ue(\langle N,<\rangle)$, while failing in $\langle N,<\rangle$.

An important general result is casting its shadows here [Goldblatt and Thomason, 1974]:

THEOREM 66. *An elementary class of Kripke frames is modally definable if and only if it is closed under the formation of generated subframes, disjoint unions and zigzag-morphic images, while its complement is closed under the formation of ultrafilter extensions.*

Proof. This argument is given in heuristic outline here, as it is one of the most elegant applications of algebraic results in modal semantics.

Evidently, modally definable classes of Kripke frames exhibit all the listed closure phenomena: the surprising direction leads from 'closure' to 'definability'.

First, notice that one closure condition can be added for free, by an earlier result. Theorem 30 implies that our class $\mathfrak{R}$ of frames is itself closed under the formation of ultrafilter extensions: if $F \in \mathfrak{R}$, then the relevant elementary equivalent $F' \in \mathfrak{R}$ ($\mathfrak{R}$ being elementary), and hence so is its zigzag-morphic image $ue(F)$.

Now the obvious strategy is to show that $\mathfrak{R}$ equals $\mathrm{MOD}(\mathrm{Th}_{\mathrm{mod}}(\mathfrak{R}))$, i.e. the class of Kripke frames verifying each modal formula which is valid throughout $\mathfrak{R}$. The nontrivial inclusion here requires us to show that

$$\text{if } F^* \models \mathrm{Th}_{\mathrm{mod}}(\mathfrak{R}), \text{ then } F^* \in \mathfrak{R}, \text{ for every Kripke frame } F^*.$$

And here is where an excursion into the realm of modal algebra will help. F^* verifies $\mathrm{Th}_{\mathrm{mod}}(\mathfrak{R})$, and hence $A(F^*)$ verifies the equational theory of the class $\{A(G) \mid G \in \mathfrak{R}\}$. (Recall the earlier correspondence between modal formulas and polynomials.) By Birkhoff's Theorem, in a suitable version, this implies that $A(F^*)$ must be constructible as a *homomorphic image* of some *subalgebra* of some *direct product* $\Pi\{A(G_i) \mid i \in I\}$, with $G_i \in \mathfrak{R}$. In a picture,

$$A(F^*) \xleftarrow[\text{homomorphism}]{\text{surjective}} \mathfrak{A} \subseteq \Pi\{A(G_i) \mid i \in I\}.$$

Now the latter algebra is isomorphic to $A(\oplus\{G_i \mid i \in I\})$, by the earlier duality. Moreover, the latter disjoint union belongs to $\mathfrak{K}$ — by the given closure conditions. So, the picture becomes, for some $G \in \mathfrak{K}$:

$$A(F^*) \xleftarrow[\text{homomorphism}]{\text{surjective}} \mathfrak{A} \subseteq A(G).$$

Now, the transformation F turns this into the corresponding row

$$FA(F^*) \xrightarrow[\text{generated subframe}]{\text{embedding as}} F(\mathfrak{A}) \xleftarrow[\text{zigzag morphism}]{\text{surjective}} FA(G).$$

But then, finally, the following walk through the diagrams suffices. $G \in \mathfrak{K} \Rightarrow FA(G) = ue(G) \in \mathfrak{K}$ (by the above observation) $\Rightarrow F(\mathfrak{A}) \in \mathfrak{K}$ (closure under zigzag images) $\Rightarrow FA(F^*) \in \mathfrak{K}$ (closure under generated subframes) $\Rightarrow F^* \in \mathfrak{K}$ ('anti-closure' under ultrafilter extensions). ∎

Actually, this result does not yet characterise **P1**, as it talks about modal definability by any set, finite or *infinite*. The additional strengthenings needed for zeroing in on **P1** are hardly enlightening, however.

The result also says a little bit more. Adding closure under ultrafilter extensions, while removing the condition of elementary definability, yields a characterisation of those classes of Kripke frames definable by means of a *canonical* modal logic in the sense of the Introduction (i.e. one which is complete with respect to its Henkin frames). Moreover, the above proof heuristics may also be used to formulate a general closure condition on classes of Kripke frames necessary and sufficient for definability by means of just *any* set of modal formulas ('SA-constructions'; cf. [Goldblatt and Thomason, 1974]).

As with the earlier ultrapower characterisation of **M1**, the above characterisation gives no *effective* information concerning the formulas in **P1**. What is needed are 'preservation theorems' giving the syntactic cash value of the given four closure conditions. Several of these have been given in [van Benthem, 1976], extending earlier results, e.g. of Feferman and Kreisel.

Here is an idea. Preservation under generated subframes allows only formulas constructed from atomic formulas and their negations, using

> $\forall, \wedge, \vee$ as well as *restricted* existential quantifiers $\exists v(Ruv\wedge$ $(u, v$ distinct).

Preservation under disjoint unions admits only one single universal quantifier in front: all others are to be *restricted* to the form $\forall v(Ruv \rightarrow)$. Finally, preservation under zigzag images forbids the negations, and we are left with

THEOREM 67. *A first-order sentence is preserved under the formation of generated subframes, disjoint unions and zigzag-morphic images if and only if it is equivalent to one of the form $\forall x\alpha(x)$, where $\alpha(x)$ has been constructed from atomic formulas using only conjunction, disjunction and restricted quantifiers.*

Proof. By elementary chain constructions, as in [Chang and Keisler, 1973, Chapter 3.1]. ■

For preservation under ultrafilter extensions, only some partial results have been found. (After all, the class of sentences preserved under such a complex operation *need* not even be effectively enumerable.)

As for the total complexity of **P1**, this may well be considerable — as was the case with **M1**. Are the two classes perhaps recursive in each other?

2.5 Modal Predicate Logic

As in much technical work in this area, modal *propositional* logic has been studied up till now. Modal *predicate* logic, however important in philosophical applications, is much less understood. (Cf. Chapter 2.5 in this *Handbook*.) Nevertheless, in the case of Correspondence Theory, an excuse for the neglect may be found in Theorem 69 below.

The unfinished state of the art shows already in the fact that no commonly accepted notion of semantic structure, or truth definition exists. Hence, we fix one particular, reasonably motivated choice as a basis for the following sketch of a predicate-logical variant of the earlier theory.

The *language* is the ordinary one of predicate logic, with added modal operators. *Structures* are tuples

$$\mathfrak{M} = \langle W, R, D, V \rangle,$$

where the *skeleton* $\langle W, R, D \rangle$ is a Kripke frame with a *domain function* D assigning sets of individuals D_w to each world $w \in W$. The valuation V supplies the interpretation of the nonlogical vocabulary at each world.

The *truth definition* explicates the notion

'$\varphi(x)$ is true in $\mathfrak{M}$ at w for d',

where the sequence d assigned to the free individual variables x comes from D_w. Its key options are embodied in the clauses for the individual quantifiers: these are to range over D_w, plus that for the modal operator:

> $\Box\varphi(x)$ is true at w for d if, for each R-alternative v for w *such that d is in* $D_v, \varphi(x)$ is true at v for d.

Thus, necessity means 'truth in all alternatives, where defined'.

As before, truth in a skeleton (at some world, for some sequence of individuals) means truth under all possible valuations. Again, in this way modal axioms start expressing properties of R, D — and their interplay.

The relevant matching 'working language' on the classical side will now be a *two-sorted* one: one sort for worlds, another for individuals. Its basic predicates are the two sortal identities, R between worlds, as well as the sort-crossing Exw : 'x is in the domain of w', or 'x *exists* at w'.

EXAMPLE 68. The Barcan Formula $\forall x\Box Ax \to \Box\forall x Ax$ defines

$$\forall wv(Rwv \to \forall x(Exv \to Exw)).$$

Proof. '$\Leftarrow$': Assume $\forall x\Box Ax$ at w, and consider any R-alternative v. For all $d \in D_v, d \in D_w$ (by the given condition), whence $\Box Ad$ holds at w — and, hence, Ad holds at v.

'$\Rightarrow$': The Barcan Formula will hold under the following particular assignment: $V_u(A, d) = 1$ if Rwu and $d \in D_w$.

This V verifies the antecedent, and hence the consequent. The relational condition follows. ∎

Thus, the Barcan Formula expresses an interaction between R and D. This is not accidental. For *pure* R-principles, we have the following *conservation* result.

THEOREM 69. *There exists an effective translation from sentences φ of modal predicate logic to formulas $p(\varphi)$ of modal propositional logic such that,*

> *if φ is equivalent to some pure $R, =$-sentence α, then $p(\varphi)$ already defines α in the sense of Section 2.2.*

Proof. p merely crosses out quantifiers in some suitable way. For full details (here and elsewhere) cf. [van Benthem, 1983]. ∎

Besides the Barcan Formula, there are three further fundamental 'de re/de dicto interchanges'. One of these provides a new example of non-first-order definability.

EXAMPLE 70.

1. $\Box\forall x Ax \to \forall x\Box Ax$ is universally valid,
2. $\exists x\Box Ax \to \Box\exists x Ax$ defines $\forall wv(Rwv \to \forall x(Exw \to Exv))$,

3. $\Box\exists x Ax \to \exists x \Box Ax$ defines an essentially higher-order condition on $R, =, E$.

Despite the superficial resemblance to the McKinsey Axiom of section 2.2., the proof for the latter result is quite different from that of Example 43. Interested readers may notice that the above principle holds in worlds with a *finite* chain of overlapping two-element successors:

$$\{1,2\}, \{2,3\}, \{3,4\}, \ldots, \{n-1, n\}, \{n, n+1\}.$$

But, it may fail in the presence of *infinite* such chains, and then compactness lurks.

Further systematic reflection on the above 'positive' result yields a method of substitutions again, with an outcome like that of Theorem 48:

THEOREM 71. *Formulas of the form $\varphi \to \psi$, with φ constructed from atomic formulas prefixed by a (possibly empty) sequence of $\forall, \Box$, using only $\wedge, \vee, \exists$ and $\Diamond$, and ψ constructed from atomic formulas using $\wedge, \vee, \exists, \Diamond$ as well as $\forall, \Box$, are all uniformly first-order definable.*

The global mathematical characterisation of first-order definability remains essentially the same in this area, whence it is omitted here.

Something which does *not* generalise easily, however, is the algebraic approach of Section 2.3. This is an endemic problem in classical (and intuitionistic) logic already: elegant algebraization stops at the gates of predicate logic. There could be an area of 'modal cylindric algebra' of course (cf. [Henkin *et al.*, 1971]), but none exists yet. (For an interesting related area, cf. the extension of modal propositional algebra to the modal program algebra of dynamic logicians (cf. [Kozen, 1979] or the Dynamic Logic chapter in volume 5 of this *Handbook*).) As a consequence, we still lack an elegant characterisation of the modally definable fragment of the present two-sorted first-order language.

What we do have, however, is such a characterisation for that same language with *parametrised predicate constants* $A(w, -)$ for the predicate constants $A(-)$ of the modal predicate logic. Thus, this is the appropriate language for the *first-order* transcription of the above truth definition. The Barcan Formula, for example, becomes

$$\begin{aligned} &\forall x(Exw \to \forall v((Ewv \wedge Exv) \to Avx)) \to \\ &\quad \to \forall v(Rwv \to \forall x(Exv \to Avx)). \end{aligned}$$

As in Theorem 18, two characteristic modal relations suffice for characterising the modal transcriptions among the class of all formulas of this language. In order to end on an optimistic note, here is the relevant result.

First, modal predicate logic knows *generated submodels*, just as in Section 2.1. Moreover, the earlier *zigzag relations* may be enriched so as to incorporate individual back-and-forth choices, as in the Ehrenfeucht–Fraïssé approach to first-order definability.

DEFINITION 72. A *zigzag connection* C between two models M_1, M_2 relates finite sequences (w, x) of equal length (w a world, x a sequence of individuals in the domain of w) in such a way that

1. all such sequences occur: those from M_1 in the domain, those from M_2 in the range of C
2. if $C(w,x)(v,y)$ and $w' \in W_1$, with $R_1ww', x \in D_w$, then $C(w',x)(v',y)$ for some $v' \in W_2$ with $R_2vv', y \in D_{v'}$,
   and analogously in the opposite direction ('world zigzag')
3. if $C(w,x)(v,y)$ and $d \in D_w$,
   then $C(w, x\frown d)(v, y\frown e)$ for some $e \in D_v$,
   and vice versa. ('individual zigzag')
4. if $C(w,x)(v,y)$, then the map $(x)_i \to (y)_i$ is a *partial isomorphism* between $\langle D_w, V_w\rangle$ and $\langle D_v, V_v\rangle$.

Now, transcriptions of modal formulas are *invariant* for generated submodels and zigzag connections, in the obvious sense. E.g. the latter have been made precisely in such a way that for modal φ,

$$\varphi \text{ is true at } w \text{ for } x \quad \text{iff} \quad \varphi \text{ is true at } v \text{ for } y, \quad \text{when } C(w,x)(v,y).$$

THEOREM 73. *A formula $\varphi = \varphi(w,x)$ of the two-sorted world/individual language is (equivalent to the transcription of) a modal formula if and only if it is invariant for generated submodels and zigzag connections.*

Proof. This follows from the main proof in [van Benthem, 1981b]. ■

On the whole, exciting technical results are yet scarce in modal predicate logic — and Correspondence Theory is no exception.

2.6 Higher-Order Correspondence

Modal formulas define second-order (Π_1^1) conditions on the alternative relation in all cases, and first-order conditions in some. In the perspective of abstract model theory, two possible generalisations arise here.

Instead of the first-order target language, one may consider suitable extensions. For instance, in Theorem 37, the relevant relational condition was definable in $L_{\omega_1,\omega}$: first-order logic with countable conjunctions and disjunctions. Not all modal formulas become definable here, however. E.g. Löb's Axiom defined a form of well-foundedness, which is known to be beyond $L_{\omega_1\omega}$, or indeed any language of the $L_{\infty\omega}$-family. On the other hand, this time for instance, the defining condition is already 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 Π^1_1. And, in fact, even the latter case itself is interesting. Which Π^1_1-sentences, for example, admit of modal definitions?

Given the general lack of semantic characterisations for such higher logics, such characterisations for their modal fragments are also difficult to obtain. One observation might be that both $L_{\omega_1\omega}$ and L^2 have the property of *invariance for partial isomorphism* (cf. van Dalen's chapter in Volume 1 of this *Handbook*). It will be of interest to study this preservation condition on modal formulas. In fact, no counter-examples have been discovered yet; but these do exist in tense logic. (The rationals $\langle Q, <\rangle$ and the reals $\langle R, <\rangle$ are a classical example of partially isomorphic structures, but there exists a tense-logical formula expressing Dedekind Completeness, which is valid on the latter, though not on the former frame.)

On the other hand, the modal propositional language could itself be strengthened, notably by the introduction of propositional quantifiers $\forall p, \exists p$, which have occurred in various places in the literature (cf. Garson's chapter in Volume 3 of this *Handbook*). Thus, e.g. $\forall p(\Box\Diamond p \to \exists q\Diamond\Box q)$ would become an admissible formula, but also $\Box\exists p\Diamond p \to \Diamond\forall q\Diamond\Box q$. Actually, there is a choice here, whether to allow the propositional quantifiers in the scope of modal operators or not. Henceforth, we consider the second, more restricted option.

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 has a reasonable motivation through the modal 'rules' mentioned in Section 3.2 below.

It has been observed by Gabbay that the following rule defines *irreflexivity* of Kripke frames:

$$\text{'if } F \vDash (\Box p \wedge \neg p) \to \varphi[w] \text{ (with } \varphi \text{ } p\text{-free), then } F \vDash \varphi[w]\text{'.}$$

The general pattern here is that of '$F \vDash \varphi[w]$ only if $F \vDash \psi[w]$', i.e. an implication of two Π^1_1-formulas, which is Δ^1_2. (It may be written either in the form $\forall\exists$ or $\exists\forall$.)

Actually, the above specific example is already Σ^1_1, as it amounts to $\forall pq((\Box p \wedge \neg p) \to q) \to \forall qq$, i.e. $\forall p((\Box p \wedge \neg p) \to \bot) \to \forall qq$, i.e. $\forall p((\Box p \wedge \neg p) \to \bot) \to \bot$, i.e. $\exists p(\Box p \wedge \neg p)$. Another relevant observation is that implications of the above form $\forall \to \forall$, if first-order definable at all, already have a first-order definable consequent. We do not go into these specific matters here, but note a general issue.

As often 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** (cf. Section 2.4), while Δ^1_2 adds conjunctions and disjunction across **P1** and the latter 'mirror image'.

QUERY. Does the second-order prenex hierarchy induce an *ascending* corresponding hierarchy of modally definable first-order principles about the alternative relation?

This possibly ascending hierarchy cannot exhaust all first-order principles, as higher-order modal formulas do retain one basic preservation property: their local truth is invariant under passing to generated subframes. (The Generation Theorem 15 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 (cf. [van Benthem, 1976, Chapter 6]). 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 of Section 2.1 are lost, however. As was observed earlier, irreflexivity $(\forall x \neg Rxx)$ becomes definable and, hence, preservation under zigzag morphisms fails. Anti-preservation under ultrafilter extensions fails, because the earlier example $\forall x\exists y(Rxy \wedge Ryy)$ becomes definable as well. (A straightforward definition uses a propositional quantifier within a modal scope: $\Diamond\forall p(\Box p \to p)$. But there is a nonembedded substitute in the form of $\exists p(\Diamond p \wedge \forall q\Box(p \to (\Box q \to q)))$.)

Thus, we arrive at the following

QUESTION. Can every almost-restricted first-order formula $\forall x\varphi(x)$ be defined at some level in the modal propositional quantifier hierarchy?

Using 'simulation' of restricted first-order quantification by propositional quantifiers, one may indeed handle most obvious cases. Here is one illustration of the procedure

EXAMPLE. Let $\varphi(x)$ be $\exists y(Rxy \wedge \forall z(Ryz \to (Rzz \vee (Rzy \wedge Rzx))))$. The idea is to define $\{x\}, \{y\}, \{z\}$, in a sense, as far as necessary (i.e. on the set consisting of x, its R-, R^2- and R^3-successors) — and then to express all desired relations between these by means of modal formulas:

> $\exists p_x(p_x \wedge\ \forall q_x(((p_x \wedge q_x) \vee \Diamond(p_x \wedge q_x) \vee \Diamond\Diamond(p_x \wedge q_x) \vee \Diamond\Diamond\Diamond(p_x \wedge q_x)) \to (\Box(p_x \to q_x) \wedge \Box\Box(p_x \to q_x) \wedge \Box\Box\Box(p_x \to q_x)))$ [this makes p_x unique to the extent indicated] $\wedge\ \exists p_y(\Diamond p_y \wedge$ [same uniqueness statement] $\wedge\ \forall p_z((\Diamond(p_y \wedge \Diamond p_z)\wedge$ [same uniqueness statement]$) \to (\forall q_z\Box\Box(p_z \to (\Box q_z \to q_z))$[i.e. '$Rzz$'] $\vee\ \Diamond\Diamond(p_z \wedge \Diamond p_x \wedge \Diamond p_y)$[i.e. '$Rzy \wedge Rzx$']$)))))$.

Accordingly, our conjecture is that the above question has a positive answer.

We conclude with one further

QUESTION. Does the addition of propositional quantifiers within modal scopes add any power of expression?

3 OTHER INTENSIONAL NOTIONS

Modal logic is only one branch, be it a paradigmatic one, of intensional logic in general. But also in other intensional areas, a Correspondence Theory is possible. In some cases, the generalisation runs smoothly: existing notions and results may be applied at once, or after only minor modification. A case in point is *tense logic*, to be treated in Section 3.1. More challenging generalisations arise when the relevant intensional semantics exhibit strong peculiarities, diverging from the earlier modal case. Sometimes, these assume the form of pre-conditions on the alternative relation; but maybe the most important hurdle is when a restriction is proclaimed on 'admissible assignments'. Both phenomena occur in *conditional logic*, the topic of Section 3.2. That, even under such circumstances, an interesting Correspondence Theory may remain, is shown by the example of *intuitionistic logic* in Section 3.3.

These two new features do not exhaust the possible semantic variation. One may also move to the interplay of different kinds of intensional operators, for instance, using correspondence to connect different alternative relations.

EXAMPLE. In *dynamic logic*, two modal operators $\Box, \Box^*$ figure, which may be provided with two alternative relations R, R^*. (Recall that $\boxed{a}$ means 'after every successful computation of a', while the intuitive meaning of $\boxed{a}^*$ is to be: 'after any finite number of runs of a'.) Now, from a correspondence point of view, the well-known *Segerberg Axioms*

$$\begin{array}{l} \Box^* p \to \Box p \\ \Box^* p \to \Box\Box^* p \\ \Box^*(p \to \Box p) \to (\Box p \to \Box^* p) \end{array}$$

define precisely the condition that

R^* coincides with the transitive closure of R.

The very exoticness of this example to many readers may help to show that Correspondence Theory is omnipresent.

No systematic developments will be given in the following sections. Their purpose is to convey an impression of notions and themes, through mainly illustrative examples. Indeed, here is where the reader may wish to carry on the torch herself.

3.1 Tense Logic

Traditionally, tense-logical structures have been taken to be temporal orders $\langle T, < \rangle$, where T consists of the points in Time, ordered by precedence $<$ ('earlier than', 'before'). The simplest formal language to be chosen has been that of Prior, adding operators G ('it is always going to be'), H ('it has always been') to some propositional base. We add F ('future'), P ('past') as derived notions. (Cf. the chapter on Basic Tense logic in volume 6 for the necessary background in tense logic.)

Of the amazing diversity of 'ontological' and 'linguistic' questions concerning this temporal semantics, only a few themes will be mentioned here. (Cf. [van Benthem, 1985] for a varied exploration.)

Explaining philosophical dicta. In his famous paper 'The Unreality of Time', the philosopher McTaggart enunciated several temporal principles. One of these reads [McTaggart, 1908]:

> "If one of the determinations past, present and future can ever be applied to (an event), then one of them has always been and always will be applicable, though of course not always the same one."

When translated into Priorean axioms, this becomes a list:

1. $Pq \to H(Fq \vee q \vee Pq)$
2. $Pq \to GPq$
3. $q \to HFq$
4. $q \to GPq$
5. $Fq \to HFq$
6. $Fq \to G(Fq \vee q \vee Pq)$.

What do these principles mean? The answer may be obtained through the method of substitutions (fitted to the temporal case — but such generalisations will be presupposed tacitly henceforth).

EXAMPLE 74.

1. defines *left-connectedness*: $\forall x \forall y < x \forall z < x(y < z \vee z < y \vee y = z)$,
2. defines *transitivity*: $\forall x \forall y < x \forall z > x\ y < z$,
3. defines $\top$,
4. defines $\top$.

If G, H had been interpreted through different relations $<_G, <_H$, then (3) and (4) would have expressed that $<_H$ is the *converse* relation of $<_G$.

5. defines *transitivity* again: $\forall x \forall y > x \forall z < x\ z < y$,

6. defines *right-connectedness*: $\forall x \forall y > x \forall z > x (y < z \vee z < y \vee y = z)$.

Thus, the McTaggart temporal picture is one of linear flow.

An incompleteness theorem. Simple transfer of earlier modal results establishes the seminal incompleteness result of [Thomason, 1972], in a very simple version.

THEOREM 75. *The tense logic axiomatised by*

$$\begin{array}{ll} H(Hp \to p) \to Hp & \textit{(Löb's Axiom)} \\ GFp \to FGp & \textit{(McKinsey Axiom)} \end{array}$$

is incomplete.

Proof. Specifically, this logic holds in no frame — and yet it is not inconsistent.

First, as to the former statement, recall from Section 2.2 that

1. Löb's Axiom defines *transitivity* of $>$ and *well-foundedness* of $<$.

By the former, $<$ is transitive as well (transitivity is 'independent of the temporal direction', or *isotropic* (cf. [van Benthem, 1985])). Thus, in this special case, Example 51 applies, and we have

2. McKinsey's Axiom defines atomicity: $\forall x \exists y > x \forall z > y\ z = y$.

A consequence of the latter property is $\forall x \exists y > x\ y < y$ (cf. Example 65(4)). So, the temporal order must contain instantaneous loops $\ldots < y < y < y < \ldots$, which contradicts well-foundedness. Therefore, our logic holds in no frame.

Nevertheless, it does hold in a *general* frame, viz. an earlier example from Section 2.1: $\langle N, <, \mathfrak{W} \rangle$, with

$$\mathfrak{W} = \{X \subseteq N \mid X \text{ is finite or } N - X \text{ is finite}\}.$$

The reason was that refutations for the McKinsey Axiom are no longer 'admissible', as these involve infinite alterations. (Thomason gives a speculation at this point concerning the Second Law of Thermodynamics: 'event patterns stabilise'.) But then, the logic cannot be inconsistent: its **K**-theorems hold in all general frames where it is valid. ■

Tense-logical axioms for the temporal order. In [van Benthem, 1985], the following fundamental axioms are derived for any temporal order induced by a *comparative* (in the linguistic sense) 'earlier than'.

1. *irreflexivity*: $\forall x\ \neg x < x$ ('no vortices in Time')

2. *transitivity*: $\forall x \forall y > x \forall z > y\ z > x$ ('flow')

3. *almost-connectedness*: ('arrows are comparative yard sticks')
   $\forall x \forall y > x \forall z\ (x < z \vee z < y)$

A version of the latter principle may also be found as the key axiom in Leibniz' relational theory of Space-Time (cf. [Winnie, 1977]).

Which tense-logical axioms correspond? From Section 2.4, we know that (1) is undefinable, (2) yields $Gp \to GGp$, while (3) just fails to fall under Theorem 67. What the latter result does give is a correspondence between

$$\forall x \forall y > x \forall z > y \forall u > x (y < u \vee u < z)$$

and

$$(F(p \wedge Fq) \wedge Fr) \to (F(p \wedge Fr) \vee F(r \wedge Fq)).$$

Another example concerns particular temporal orders. One can never hope to fully define such frames categorically by their tense-logical theories. For, by the Generation Theorem, tense-logical formulas cannot distinguish between one single, or several parallel flows of Time — which latter picture is so familiar from contemporary science fiction. Still, if *disjoint unions* of frames are disregarded, we have

THEOREM 76. *$\langle N, < \rangle$ is defined categorically by the axioms*

$$\begin{array}{l} H(Hp \to p) \to Hp \\ Pp \to H(Fp \vee p \vee Pp) \\ Fp \to G(Fp \vee p \vee Pp) \\ FT \\ G(Gp \to p) \to (FGp \to Gp) \end{array}$$

The proof is omitted here.

But, e.g. the integers $\langle Z, < \rangle$ cannot be thus defined; as the contraction to a single point remains a zigzag morphism preserving their theory. ($\langle N, < \rangle$ was unafflicted this time: in tense logic, zigzag morphism have *two* backward relational clauses — whence, the earlier contraction fails to quality.)

Time and modality. Combined modal-tense logics with two alternative relations $R, <$ have been repeatedly proposed. For instance, in [White, 1981] we find a logic with characteristic axioms

$$\begin{array}{ll} Gp \to GGp, Fp \to G(Fp \vee p \vee Pp), PT & \text{(D4.3)} \\ Pq \to \Box Pq & \text{('irrevocable past').} \end{array}$$

This logic is claimed to be appropriate for an analysis of the famous Diodorean 'Master Argument', identifying *possibility* with *actual or future truth* — a version of what was later to become known as the principle of Plenitude: all metaphysical possibilities are eventually realised in this World.

Our analysis of this claim runs as follows. $Gp \to GGp$ defines *transitivity* for $<$, the McTaggart Axiom defines right-connectedness; while PT defines *left-succession*: $\forall x \exists y\ y < x$. The additional 'mixing postulate' defines

$$\forall xy(Rxy \to \forall z(z < x \to z < y)).$$

CLAIM (1). $\forall xy(Rxy \to (y < x \vee y = x \vee x < y))$.

Proof. Assume Rxy. Let $z < x$ (by left-succession). Then $z < y$ ('mix'). The conclusion follows by right-connectedness. ■

CLAIM (2). $\forall xy(Rxy \to (x < y \vee x = y))$.

Proof. If Rxy and $y < x$, then $y < y$ ('mix'): *contra* irreflexivity. ■

The outcome is this: *without* ever using transitivity, but *with* irreflexivity (which is presupposed in White's whole set-up), a relational condition follows which is indeed defined by the Diodorean challenge:

$$\Diamond p \to (Fp \vee p).$$

This is only one of the many possible semantics for temporal modalities, of course. The correspondence aspect of, e.g. the Occamist 'branching time' of [Burgess, 1979] remains to be explored.

Alternative temporal ontologies. Recently 'interval structures' have been proposed as an alternative for the above traditional point ontology. From the manifesto of [Humberstone, 1979], a picture emerges of triples

$$\langle I, \subseteq, < \rangle,$$

where $\subseteq$ is *inclusion* among intervals, and $<$ total *precedence*.

Here again, correspondences prove useful in exploring proposed principles. The language has the ordinary tense-logic operators, as well as a modality $\Box$ ('in all subintervals'). In this notation, Humberstone's base logic has for its basic axioms

1. $Fp \to \Box Fp$
2. $F\Diamond p \to Fp$
3. $\Diamond F\Box p \to (\Diamond p \vee Fp)$.

By the earlier method of substitutions, equivalents may be found illuminating these:

1. defines $\forall x \forall y > x \forall z \subseteq x\ y > z$,

a property known as *left monotonicity*,

2. defines $\forall x \forall y > x \forall z \subseteq y\ z > z$,

its dual property of *right monotonicity*. Finally,

3. defines $\forall x \forall y \subseteq x \forall z > y\ (\exists u \subseteq z : u \subseteq x \vee \exists u \subseteq z : u > x)$,

a form of a principle known as *convexity*. ('Stretches of time should be uninterrupted'.)

Starting from the other side, one may impose basic postulates on $\subseteq, <$, asking for definitions in this 'interval tense logic'. For $<$, these might be the earlier-mentioned ones, for $\subseteq$, a minimum seems to be the requirement of *partial order*, while monotonicity (and convexity) take care of minimal connections between $<, \subseteq$. This would add only two axioms to the preceding ones, viz. **S4** for inclusion. The further condition of *anti-symmetry* is not definable — as may be seen by noting that the map $n \mapsto n$ (modulo 2) is a $\subseteq$-zigzag morphism sending the anti-symmetric frame $\langle Z, \leq \rangle$ to the non-antisymmetric one $\langle \{0,1\}, \{\langle 0,0\rangle, \langle 0,1\rangle, \langle 1,0\rangle, \langle 1,1\rangle\}\rangle$.

Many more examples of further correspondences on top of this foundation may be found in Chapter II.3.2 of [van Benthem, 1985].

3.2 Conditionals

From among the teeming multitude of 'conditional logics', three specimens have been included here. As no work of the present kind has been done in this area at all, the following considerations are still very much first steps. (Cf. the Conditional Logic chapter in volume 5 for a discussion of conditional logics.)

Constructive implication

Perhaps the single most effective argument in favour of constructive, as opposed to classical implication is the natural deduction analysis. The *natural* rules for $\to$-introduction and $\to$-elimination give us only a fragment of all classical pure $\to$-tautologies; axiomatised by

(A1) $\varphi \to (\psi \to \varphi)$

(A2) $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$

plus the rule of *modus ponens*. A principle notably outside of this class is Peirce's Law

$$((\varphi \to \psi) \to \varphi) \to \varphi.$$

But really, the same elegance shows up in the Henkin completeness proof. In the usual proof, one starts from a given consistent set — and then has to extend this arbitrarily to just *any* maximally consistent one, in order to 'break down' implications according to the classical truth table. A *canonical* model construction rather uses a unique natural model, viz. that consistent set together with all its consistent extensions, exploiting the evident decomposition rule

$$\Sigma \vdash \varphi \to \psi \text{ if and only if } \forall \Sigma' \supseteq \Sigma : \textit{ if } \Sigma' \vdash \varphi, \textit{ then } \Sigma' \vdash \psi.$$

A perfect match arises with the following semantics. Structures are general frames $F = \langle W, R, \mathfrak{W} \rangle$, where R corresponds to the above inclusion relation, and $\mathfrak{W}$ consists of all *R-hereditary* sets of worlds. (Propositions represent R-cumulative knowledge on this view.)

A direct study of the above logic on these frames would yield rather clumsy conditions. One case will be exhibited, as it illustrates a variant concept of correspondence, viz. *correspondence for rules* rather than *axioms*.

EXAMPLE 77. Modus Ponens defines the condition 'every world belongs to some finite R-loop'.

Proof. '$\Leftarrow$': Suppose that $xRx_1R \ldots Rx_nRx$. Let $V(p), V(q)$ be R-hereditary subsets of W, such that $p, p \to q$ hold at x. Then, successively, p, q hold at $x_1, \ldots, x_n$, and finally at x.

'$\Rightarrow$': Suppose that x belongs to no finite R-loop. Set $V(p) :=$ the smallest R-hereditary set containing $x, V(q) =$ the R-hereditary closure of $\{y \mid Rxy\}$. This verifies $p, p \to q$ at x; without verifying q. ■

What will be done instead is to postulate the *partial order* behaviour of $\subseteq$: *reflexivity, transitivity* and *antisymmetry*. Finer peculiarities of (A1), (A2) remain undetectable below this threshold.

Further restrictions on R may now be imposed by stronger axioms; e.g. we can see why Peirce's Law is characteristic for classical logic.

EXAMPLE 78. Peirce's Law defines the restriction to single points:

$$\forall xy(Rxy \to y = x).$$

Proof. '$\Leftarrow$': A simple calculation suffices.

'$\Rightarrow$': Suppose that $Rxy, x \neq y$. Set $V(q) = \varnothing, V(p) = \{z \mid Rxz \wedge x \neq z\}$. This makes $(p \to q) \to p$ true at x (notice that $p \to q$ is false at x itself), while falsifying p. (By the way, that V is admissible, i.e. that $V(p)$ is R-hereditary, follows from the above general assumption.) ■

But 'intermediate' implication axioms exist as well.

EXAMPLE 79. The following principle

$$((p \to q) \to p) \to (((q \to r) \to q) \to p)$$

defines a maximal length 3 for R-chains:

$$\forall xy(Rxy \to \forall z(Ryz \to (x = y \vee y = z \vee \forall u(Rzu \to z = u)))).$$

Proof. Here is the relevant counter-example for the argument in the '$\Rightarrow$'-direction. Assume that $xRyRzRu$, while $x \neq y, y \neq z, z \neq u$. Set $V(r) = \emptyset, V(q) = \{v \mid Ruv \wedge u \neq v\} \cup \{v \mid Ryv \wedge \neg Rvz\}, V(p) = \{v \mid Ryv \wedge y \neq v\}$. The principle will be falsified at y. ■

It has not been possible to find other types of intermediate example. Hence, we conclude with a

CONJECTURE. All principles of pure constructive implication define first-order constraints on R; viz. restrictions to some finite chain length.

Relevant implication

Of the various proposed semantics for relevance logic, here is a perspicuous example from [Gabbay, 1976, Chapter 15]. Structures are now tuples $\langle W, R, V, 0\rangle$, where 0 is a special world providing a vantage point from which to compare other worlds through the *ternary* relation R. Intuitively, R_abc is to mean that b is 'included' in c, at least from the perspective of a. (One might think of, for example, 'a-local inclusion': $a \cap b \subseteq a \cap c$.) No prior conditions are imposed on this relation.

This is not to say that these are not to be found at all. For instance, it may be shown that the mentioned local inclusion relation is characterised by two *betweenness* axioms:

1. $R_abc \leftrightarrow R_bac$ (interchanging boundaries)
2. $(R_abc \wedge R_dae \wedge R_dbe) \to R_dce$

(I.e. if $c \in [a, b], a \in [d, e], b \in [d, e]$, then $c \in [d, e]$: a form of convexity.)

The explication of implication reads as follows:

> $\varphi \to \psi$ is true at a iff, for all b, c such that R_abc, if φ is true at b, then ψ is true at c.

As it stands, this definition makes *no* implication laws universally valid. To obtain at least some indubitable principle, one therefore imposes a restriction on valuations. The most urgent case is that of $p \to p$. On the above bare semantics, it would correspond to $\forall xyz(R_xyz \to y = z)$, collapsing the ternary relation. To avoid this, one again requires 'cumulation':

> valuations V are only to assign subsets X of W subject to the constraint that $\forall xy \in W(R_0xy \to (x \in X \to y \in X))$.

If this constraint is to extend automatically to sets X defined by complex implicational formulas, then a mild form of *transitivity* is to be imposed on the ternary relation after all:

$$\forall xyzu((R_0xy \wedge R_yzu) \rightarrow R_xzu).$$

Notice how this relates perspectives from different vantage points.

But then, if reasonable forms of transitivity have become respectable, we also add $(*)\forall xyzu((R_0xy \wedge R_0yz) \rightarrow R_0xz)$.

Now, at last, some genuine correspondences arise — of a 'local' sort (cf. Section 2.2).

EXAMPLE 80.

1. Modus Ponens defines R_000,
2. Axiom A1 defines a curious form of 'transitivity':
   $\forall xyzu((R_0xy \wedge R_yzu) \rightarrow R_0xu)$.

Proof. (Case (1) only) '$\Leftarrow$': This direction is immediate.

'$\Rightarrow$': Let $V(p) = \{0\} \cup \{x \mid R_00x\}, V(q) = \{x \mid R_00x\}$. By the above principle $(*)$, both assignments are admissible. Clearly, both p and $p \rightarrow q$ are true at 0, whence also q: i.e. R_000. ■

Obviously, the second principle is not very plausible — but then, neither is (A1) for a relevance logician.

A more interesting phenomenon in relevance logic, from the present point of view, is the treatment of *negation*. This formerly inconspicuous notion is now interpreted using a '*reversal operation*' $^+$ on worlds:

$$\neg\varphi \text{ is true at } a \quad \text{iff} \quad \varphi \text{ is true at } a^+.$$

In this light, new combined correspondences appear, such as that between Contraposition and the reversal law

$$\forall xy(R_0xy \rightarrow R_0y^+x^+).$$

Correspondence Theory may be applied to any kind of semantic entity.

Counterfactual implication

Ramsey told us to evaluate conditionals as follows. Make the minimal adjustment of your stock of beliefs needed to accommodate the antecedent: then see if the consequent follows. Various syntactic and semantic implementations of this view exist, of which that of [Lewis, 1973] has deservedly won the greatest favour. A counterfactual $\varphi \,\Box\!\!\rightarrow \psi$ is true in a world, on his account, if ψ is true in all worlds most similar to that world *given that* φ holds in them.

As the preceding account has some difficulties in the infinite case, let us consider *finite* models $\langle W, C, V\rangle$, where C is a *ternary* relation of comparative similarity:

$$C_x yz \text{ for: `}y \text{ is closer to } x \text{ than } z \text{ is'}.$$

Lewis gives three basic conditions on the relation 'no closer':

1. *transitivity*: $\forall xyzu((\neg C_x yz \wedge \neg C_x zu) \rightarrow \neg C_x yu)$,
2. *connectedness*: $\forall xyz(\neg C_x yz \vee \neg C_x zy)$,
3. *egocentrism*: $\forall xy(\neg C_x xy \rightarrow x = y)$.

Rewriting these for 'closer', one finds to one's surprise that (2) is rather weak, being merely

$2'$. *asymmetry*: $\forall xyz(C_x yz \rightarrow \neg C_x zy)$.

On the other hand, (1) becomes a strong principle

$1'$. $\forall xyu(C_x yu \rightarrow \forall z(C_x yz \vee C_x zu))$,

which we knew as *almost-connectedness* back in Section 3.1.

From asymmetry and almost-connectedness, one may derive ordinary *transitivity* and *irreflexivity*, whence the three 'comparative' axioms of Section 3.1 emerge. These principles justify the appealing picture of 'similarity spheres' around the reference world x.

The tendency has been since 1973 to retain only *transitivity* and *irreflexivity* as fundamental pre-conditions on C, leaving various forms of connectedness as optional extras. Thus, one finds an axiomatisation of this austere minimal conditional logic in [Burgess, 1981].

The truth definition in this case may be taken to be the following:

$$\varphi \mathrel{\Box\!\!\rightarrow} \psi \text{ is true at } w \text{ if } w \text{ holds in all } \varphi\text{-worlds } C\text{-closest to } w.$$

Indeed, this clause verifies the following list of principles without further ado:

$$\begin{array}{l} p \mathrel{\Box\!\!\rightarrow} p, \\ p \mathrel{\Box\!\!\rightarrow} q, p \mathrel{\Box\!\!\rightarrow} r \vdash p \mathrel{\Box\!\!\rightarrow} q \wedge r, \\ p \wedge q \mathrel{\Box\!\!\rightarrow} p, \\ p \mathrel{\Box\!\!\rightarrow} r, q \mathrel{\Box\!\!\rightarrow} r \vdash p \vee q \mathrel{\Box\!\!\rightarrow} r. \end{array}$$

It is only the last one which requires *transitivity*:

$$p \mathrel{\Box\!\!\rightarrow} q \wedge r \vdash p \wedge q \mathrel{\Box\!\!\rightarrow} r.$$

Egocentrism is restored by adding the principle of Modus Ponens:

$$p \mathrel{\Box\!\!\rightarrow} q, p \vdash q$$

But, the original Lewis logic contained even further principles, such as the formidable

$$((p \vee q) \mathrel{\Box\!\!\rightarrow} p) \vee \neg((p \vee q) \mathrel{\Box\!\!\rightarrow} r) \vee q \mathrel{\Box\!\!\rightarrow} r.$$

What does it express? As it happens, it restores *almost-connectedness*.

Proof. First, the axiom is valid under this additional assumption — by the above discussion.

Next, suppose almost-connectedness fails; i.e. for some $xyzu$ we have: $C_xyz, \neg C_xyu, \neg C_xuz$. By transitivity, it follows that $\neg C_xzu$. Now, set $V(p) = \{y\}, V(q) = \{z, u\}, V(r) = \{y, u\}$. Then z is q-closest among the worlds falsifying r. The two $p \vee q$-closest worlds y, u both verify r. Finally, p fails in the $p \vee q$-closest world u. Thus, Lewis' axiom has been refuted. ■

Finally, to mention an example outside of Lewis' original logic, there is the Stalnaker principle of 'Conditional Excluded Middle':

$$p \mathrel{\Box\!\!\rightarrow} q \wedge p \mathrel{\Box\!\!\rightarrow} \neg q.$$

As was stated in the Introduction, this axiom even requires the similarity order to be a *linear* one. In the present finite case, this means that the above truth definition reduces to:

$\varphi \mathrel{\Box\!\!\rightarrow} \psi$ is true at w if ψ holds in the closest φ-alternative to w.

And that was the original Stalnaker explication of conditionals.

The previous examples were all conditional axioms without nestings of $\mathrel{\Box\!\!\rightarrow}$. This is typical for most current logics in this area. Relational conditions matching these have invariably been found to be first-order ones. Hence, in view of Theorem 38, here is our

CONJECTURE. All counterfactual axioms without nestings of conditionals are first-order definable.

The reason for this restriction lies in the motivation for the present area. Entailment conditionals such as constructive implication, or modal entailment have often been proposed out of dissatisfaction with classical 'nested principles', such as, say, $p \to (q \to p)$ or Peirce's Law. The non-nested classical fragment was not called into question. Counterfactual conditionals, however, typically disobey classical implicational logic at the level of non-nested inferences, such as the monotonicity rule from $p \to q$ to $p \wedge r \to q$.

Nevertheless, there are intrinsic reasons to be found inside the above semantics for considering nested axioms after all. For, one obvious omission in the above list of semantic conditions was the lack of *index principles* relating the perspectives of different worlds. For instance, when we read C for a moment as relative proximity in Euclidean space, we find the following *Triangle Inequality*

$$\forall xyz((C_x yz \wedge C_z xy) \rightarrow C_y xz).$$

And there are other elegant principles of this kind.

Now, it is easily seen that such index principles are just what is involved when nested counterfactuals are evaluated: the perspective starts shifting. Thus, it will be rewarding to have correspondences here as well. One, not too exciting example is the following. The Absorption Law

$$p \mathrel{\Box\!\!\rightarrow} (q \mathrel{\Box\!\!\rightarrow} r) \vdash (p \wedge q) \mathrel{\Box\!\!\rightarrow} r$$

defines the index principle

$$\forall xyz(C_x yz \rightarrow \forall u \neg C_y uz).$$

Better examples are still to be found. Indeed, e.g. the counterfactual logic of Euclidean space, the most natural geometric representation of our similarity pictures, is still a mystery.

3.3 Intuitionistic Logic

Constructive conditional logic is only a part of the full intuitionistic logic, whose Kripke semantics extends the earlier constructive models. In this section, a sketch will be given of an Intuitionistic Correspondence Theory. (For details on intuitionistic logic, cf. van Dalen's chapter in volume 7 of this *Handbook*.)

Kripke semantics, intermediate axioms and correspondence.

DEFINITION 81. An *intuitionistic Kripke model* M is a tuple $\langle W, \subseteq, V\rangle$, where $\subseteq$ is a partial order ('possible growth') on W ('stages of knowledge'). The valuation V assigns $\subseteq$-closed subsets of W to proposition letters ('cumulation of knowledge').

The truth definition has the following familiar pattern,

$M \nvDash \bot[w]$	for all $w \in W$,
$M \vDash \varphi \rightarrow \psi[w]$	if $M \vDash \psi[v]$ for all $v \supseteq w$ such that $M \vDash \varphi[v]$,
$M \vDash \varphi \wedge \psi[w]$	if $M \vDash \varphi[w]$ and $M \vDash \psi[w]$,
$M \vDash \varphi \vee \psi[w]$	if $M \vDash \varphi[w]$ or $M \vDash \psi[w]$.

Negation is defined as usual ($\neg\varphi$ becoming $\varphi \rightarrow \bot$).

The pre-condition of partial order was motivated earlier on. But, other choices may be defended as well. As is well-known, the above semantics was derived from the modal one, through the *Gödel translation* g:

$$
\begin{aligned}
&g(p) = \Box p\\
&g(\varphi \to \psi) = \Box(g(\varphi) \to g(\psi))\\
&g(\varphi \wedge \psi) = g(\varphi) \wedge g(\psi)\\
&g(\varphi \vee \psi) = g(\varphi) \vee g(\psi)\\
&g(\bot) = \bot.
\end{aligned}
$$

Now, there is a whole range of modal logics whose 'intuitionistic fragment' (through g) coincides with intuitionistic propositional logics. Amongst others, we have the

THEOREM 82. *Let X be any modal logic in the range from* **S4** *to* **S4.Grz** $=$ **S4** *plus the Grzegorczyk Axiom*

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

Then, for all intuitionistic formulas φ, φ is intuitionistically provable in Heyting's logic if and only if $g(\varphi)$ is a theorem of X.

The earlier modal correspondences yield a corresponding semantic range, between 'pre-orders' (*reflexive* and *transitive*) and 'trees':

EXAMPLE 83. Grzegorczyk's Axiom defines the combination of (i) reflexivity, (ii) transitivity, and (iii) well-foundedness in the following sense: 'from no w is there an ascending chain $w = w_1 \subseteq w_2 \subseteq \dots$ with $w_i \neq w_{i+1} (i = 1, 2, \dots)$'.

Proof. This goes more or less like the closely related Axiom of Löb. By the way, notice that (iii) implies anti-symmetry. Note also that, semantically, Grzegorczyk's axiom alone implies the **S4**-laws: syntactic derivations to match were found around 1979 by W. J. Blok and E. Pledger. ■

Thus, a case may also be made for the *Tree of Knowledge* as a basis for intuitionistic semantics. Nevertheless, we shall stick to partial orders for a start.

Above **S4Grz**, modal logics start producing greater g-fragments — the so-called *intermediate logics*, ascending to full classical logic. Intermediate axioms impose various restrictions on the pattern of growth for knowledge, classical logic forcing the existence of single ('complete') nodes.

EXAMPLE 84. (i) Excluded Middle $p \vee \neg p$ defines $\forall x \forall y (x \subseteq y \to x = y)$.

Proof.'$\Leftarrow$' is immediate.

'$\Rightarrow$': Suppose $x \subseteq y, x \neq y$. (By anti-symmetry then $y \not\subseteq x$.) Set $V(p) = \{z \mid y \subseteq z\}$. This falsifies both p and $\neg p$ at x. ■

(ii) *Weak Excluded Middle* $\neg p \vee \neg\neg p$ defines directedness.

Proof. '$\Leftarrow$': Suppose that $\neg p$ fails at x; say p holds at $y \supseteq x$. Then consider any $z \supseteq x$. As it shares a common successor with y, *and* $V(p)$ *is* $\subseteq$*-hereditary*, it has a successor verifying p, whence $\neg p$ fails at z. So $\neg\neg p$ holds at x.

'$\Rightarrow$': Suppose that $x \subseteq y, z$, where y, z share no common successors. Set $V(p) = \{u \mid z \subseteq u\}$. (Like above, this is a $\subseteq$-closed set.) Notice that $x, y \notin V(p)$. It follows that $\neg p$ fails at x (consider z), but $\neg\neg p$ fails as well (consider y). ■

(iii) *Conditional Choice* $(p \to q) \vee (q \to p)$ *defines connectedness.*

Proof. '$\Leftarrow$': Suppose that $p \to q$ fails at x; i.e. some $y \supseteq x$ has p true, but q false. Now consider any $z \supseteq x$ such that q holds. Either $z \subseteq x$, but then, by $\subseteq$-heredity, q is true at y (*quod non*), or $y \subseteq z$, and so, again by $\subseteq$-heredity, p is true at z, i.e. $q \to p$ is true at x.

'$\Rightarrow$': Let $x \subseteq y, z$ with $y \not\subseteq z, z \not\subseteq y$. Set $V(p) = \{u \mid y \subseteq u\}, V(q) = \{u \mid z \subseteq u\}$. Then $p \to q$ fails at x (watch y), and $q \to p$ fails as well (watch z). ■

Much more forbidding principles than these have been proposed as intermediate axioms. But surprisingly, these usually turned out to be first-order definable:

EXAMPLE 85. (i) The Stability Principle $(\neg\neg p \to p) \to (p \vee \neg p)$ defines

$$\begin{aligned}&\forall x \neg\exists yz\ (x \subseteq y \wedge x \subseteq z \wedge \neg\exists u(y \subseteq u \wedge z \subseteq u) \wedge \\ &\quad \wedge \forall u(\forall s(u \subseteq s \to \exists t(s \subseteq t \wedge z \subseteq t)) \to \neg\exists v(u \subseteq v \wedge y \subseteq v))).\end{aligned}$$

(ii) The Kreisel-Putnam Axiom $(\neg p \to (q \vee r)) \to ((\neg p \to q) \vee (\neg p \to r))$ defines

$$\begin{aligned}&\forall x \neg\exists yz\ (x \subseteq y \wedge x \subseteq z \wedge \neg y \subseteq z \wedge \neg z \subseteq y \wedge \\ &\quad \wedge \forall u((x \subseteq u \wedge u \subseteq y \wedge u \subseteq z) \to \exists v(u \subseteq v \wedge \neg y \subseteq v \wedge \neg z \subseteq v))).\end{aligned}$$

No matter how complex such axioms may seem at first sight, proofs of the above assertions are quite simple exercises in 'imagining what a counterexample would look like'.

This recurrent experience led to the following *conjecture* in [van Benthem, 1976]:

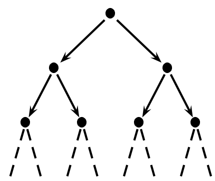
> All intermediate axioms express first-order constraints on growth of knowledge.

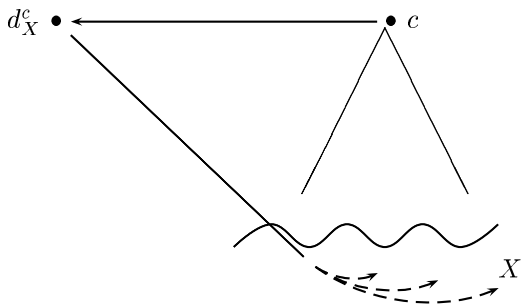
Two conjectures refuted. The earlier hope was all but given up in the first version of this chapter; as 'Scott's Rule' turned out to be an essentially higher-order intermediate inference. The relevant argument was sharpened somewhat by P. Rodenburg:

THEOREM 86. *Scott's Axiom* $((\neg\neg p \to p) \to (p \vee \neg p)) \to (\neg p \vee \neg\neg p)$ *defines no first-order condition on partial orders.*

Proof. An elaborate Löwenheim–Skolem argument works, in the spirit of Example 43. As an illustration of the non-triviality of our present subject matter, it follows here.

Step 1: Consider the following Kripke frame $\langle W, \subseteq \rangle$:





W consists of the *infinite binary tree* T, together with, for each node c in T and each $\subseteq$-*hereditary, cofinal* set X in T_c (i.e. the subtree with root c), some point d^c_X. $\subseteq$ is the usual order on T, together with

- $c \subseteq d^c_X \subseteq x$, for all $x \in X$
- $d^c_X \subseteq d^c_{X'}$, if $X' \subseteq X$.

CLAIM. *Scott's Axiom is true in* $\langle W, \subseteq \rangle$.

PROOF. First, let $c \in T$ be a putative refutation. I.e., for some valuation V,

1. $(\neg\neg p \to p) \to p \vee \neg p$ is true at c,
2. $\neg p \vee \neg\neg p$ is false at c.

Then consider the node d^c_X , where X is the cofinal hereditary set

$$T_c \cap (V(p) \cup V(\neg p)).$$

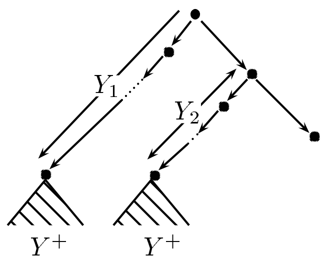
One verifies successively that $\neg\neg p \to p$ is true at d^c_X, whereas both $p, \neg p$ are false. (E.g. if p were true at d^c_X, then p is true throughout X, whence $\neg\neg p$ is true at c — whereas (2) says the opposite.) Thus, we have a contradiction with (1).

A similar argument works for the case where c is of the form d^c_X itself. ∎

Step 2: A matter of cardinality:

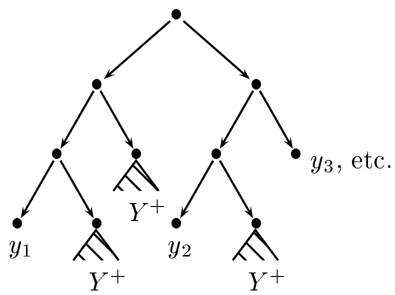
CLAIM. *The above Kripke frame is uncountable.*

PROOF. In particular, there are $2^{\aleph_0}$ nodes of the form d^c_X. For, each subset Y of N may be coded as follows, using (distinct) hereditary cofinal subsets Y^+ of the infinite binary tree. Let $Y = \{y_1, y_2, y_3, \ldots\}$.



etc. going down the extreme right branch using the extreme left branches to code $y_1, y_2, y_3, \ldots$.

For all nodes not arrived at in this way, one makes Y^+ cofinal by means of the following stipulation:

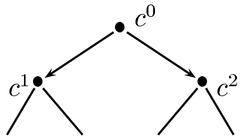


∎

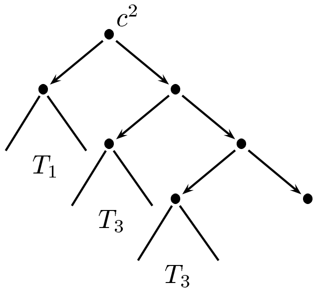
Step 3: Take any *countable* elementary substructure F of $\langle W, \subseteq\rangle$ containing the original binary tree.

CLAIM. *Scott's Axiom may be falsified in F.*

PROOF. Consider T as a double tree



and again T_{c^2} a countable sequence of 'trees on a string':



Let $D_{X_1}, D_{X_2}, \ldots$ be an enumeration of the points $d_X^{C_0}$ remaining in F. Notice that, for each $i \in N$,

1. finite intersections $T_i \cap X_1 \cap \ldots \cap X_n$ are still hereditary cofinal in T_i,
2. the total intersection $T_i \cap \{X_j \mid j = 1, 2, \ldots\}$ is empty.

As for the latter observation, it suffices to see that the assertion

$$\forall x \exists d_X^{C_0} \text{ with } d^{C_0} \not\subseteq x,$$

which holds in $\langle W, \subseteq\rangle$, can be expressed in first-order terms in $\langle W, \subseteq\rangle$; and, hence, it has remained valid in the elementary substructure F.

Now, define

$$\begin{aligned} &X_1^* = X_1 \\ &X_{n+1}^* = X_1 \cap \ldots X_k \text{ for the smallest } k \text{ such that } T_{n+1} \cap X_1 \cap \\ &\quad \ldots \cap X_k \subsetneqq T_{n+1} \cap X_n^*. \end{aligned}$$

Scott's Axiom may now be falsified at c^0, by setting

$$X^* = \cup \{T_i \cap X_i^* \mid i = 1, 2, \ldots\}, V(p) = \{y \mid \exists x \subseteq y \; x \in X^*\}.$$

to see this, notice, that successively,

1. each point d_{X_i} has a successor (in T_i) outside of $V(p)$,
2. $(\neg\neg p \to p) \to p \vee \neg p$ holds at c^0,

3. $\neg p \vee \neg\neg p$ fails at c^0. ■

We conclude that Scott's Axiom is not first-order definable — not being preserved under elementary subframes. ■

This complex behaviour disappears on better-behaved structures.

OBSERVATION 87. On *trees*, Scott's Axiom defines the first-order condition

$$\forall x \neg \exists yzu\,(x \subseteq y \wedge x \subseteq z \wedge z \subseteq u \wedge z \neq u \wedge \neg \exists v(y \subseteq v \wedge z \subseteq v)).$$

This, and other experiences of its kind, led to a revised guess in the first version of this chapter: On trees, all intermediate axioms express first-order constraints on descendance. A proof sketch was added, involving *semantic tableaux* as 'patterns of falsification', to be realised in trees.

This conjecture was 'almost' refuted in [Rodenburg, 1982]. The semantic tableau method runs into problems with *disjunctions*, and indeed we have the following counter-example.

EXAMPLE 88. Consider the formula

$$\phi = ((\neg p \wedge \neg q \wedge \neg r) \to (p \wedge q \wedge r)) \to (\neg p \wedge \neg q \wedge \neg r)$$

with the simultaneous substitution of: $p\&q$ for p, $p\&\neg q$ for q, and $\neg p\&q$ for r. This ϕ is not first-order definable on partial orders. On suitably tree-like structures, it expresses the lack of '3-forks' of immediate successors as well as the absence of infinite comb-like structures.

On trees, this negative example probably still works — but there is an instructive difficulty here. The class of trees itself has a higher-order definition; Π^1_1, to be precise. Therefore, current model-theoretic arguments for disproving first-order definability (compactness, Löwenheim–Skolem) run the risk of employing constructions leading outside of this class. Higher-order preconditions are a problem for our Correspondence Theory.

To illustrate this from a purely classical angle, the reader may consider a related problem, showing how soon the familiar methods of model theory fail us. *Finiteness* is first-order undefinable on *partial orders*, even on *trees*. It is thus definable on *linear trees*, however, viz. by 'every non-initial node has an immediate predecessor'. What about the (at most) binary trees? This intermediate case seems to be open.

The state of the subject. The progress of science is sometimes startling. Where the first version of this chapter (1981) had some tentative examples, enlightenment reigns in the report [Rodenburg, 1982]. Of its many topics, only a few will be mentioned here.

First, there are several semantic options — as indicated above, ranging from *partial orders* via '*downward linear orders*' to *trees*. But moreover, there is a legitimate choice of language. Despite appearances, it is the *disjunction* clause which is now strongly constructive in intuitionistic Kripke semantics. ('Choose *now*!' Classical logic would have a more humane clause in this setting: $\Box\Diamond(\varphi \vee \psi)$, i.e. '$\varphi$ or ψ eventually'.) Thus, it is of interest to consider both the full language and its $\vee$-free fragment.

The semantic tableau method mentioned above, in combination with the above counter-examples, has led to the results in the following scheme:

All formulas first-order definable	Partial orders	Downward linear orders	Trees
without $\vee$	YES	YES	YES
with $\vee$	NO	NO	?

But there are also matters of 'fine structure'. For instance, Scott's Axiom had only one proposition letter — and for such intuitionistic formulas we have the beautiful Rieger–Nishimura lattice. Now, Scott's Axiom merely seemed a fit candidate for a counter-example among the intermediate axioms existing in the literature. Rodenburg has proved that it is also *minimal* in the Rieger-Nishimura lattice with respect to non first-order definability. (More precisely, an intuitionistic formula with one proposition letter is first-order definable on the partial orders if and only if it is equivalent to one of $A_1, \ldots, A_9$ in the lattice.)

In the counter-examples needed for the latter result, a uniform method may be seen at work: *compactness*, in the form that sets of formulas which are finitely satisfiable in *finite* models are also simultaneously satisfiable (in some *infinite* model). Now, indeed, intuitionistic truth has a close connection with truth in finite submodels (cf. [Smoryński, 1973]). Our question is whether this may lead to the following improvement in the mathematical characterisation of first-order definability as given in Section 2.2.

CONJECTURE. An intuitionistic formula φ is first-order definable if and only if φ is preserved under *ultraproducts of finite frames*.

Intuitionistic definability. As with the direction 'from intensional to classical', the case 'from classical to intuitionistic definability' shows many resemblances with our earlier modal study. For instance, a Goldblatt–Thomason type characterisation was proved in [van Benthem, 1983] (cf. our earlier Theorem 66):

> A first-order constraint on the growth pattern is intuitionistically definable if and only if it is preserved under the formation of *generated subframes, disjoint unions, zigzag-morphic images, filter extensions* and '*filter inversions*'.

Merely in order to illustrate this topic, which has a wider semantic significance, here is a sketch of the representation theory in the background.

On the algebraic side, the intuitionistic language may be interpreted in *Heyting Algebras* $\langle A, 0, 1, +, \cdot, \Rightarrow \rangle$ satisfying suitable postulates. Now, each *Kripke* (*general*) *frame* in the above sense induces such a Heyting Algebra, through its $\subseteq$-*hereditary* sets, provided with suitable, obvious operations. But also conversely, a *filter representation* now takes Heyting Algebras to Kripke general frames. Indeed, the earlier categorial duality (cf. Section 2.3) is again forthcoming.

The more general semantic interest of the construction is this. Despite the superficial similarity with structures consisting of the 'complete' possible worlds, intuitionistic Kripke models should be regarded as patterns of stages of *partial* information. This comes out quite nicely in the above representation, where 'worlds' are no longer complete *ultra*filters, but merely filters (in the $\vee$-free case) or 'splitting' filters (for the full language). Filters F merely satisfy the closure condition that

$$a, b \in F \text{ iff } a \cdot b \in F,$$

a minimal requirement on partial information. Also quite suggestively, the 'modal' alternative relation collapses into inclusion ('growth'):

$$\forall a \Rightarrow b \in F \, \forall a \in F' : b \in F' \quad \text{iff} \quad F \subseteq F'.$$

The present-day supporters of 'partial models' and 'information semantics' would do well to study intuitionistic logic.

Predicate logic. Again, correspondence phenomena do not stop at the frontier of predicate logic. This will be illustrated by means of some intuitionistic examples.

Kripke models $M = \langle W, \subseteq, D, V \rangle$ will now be of the usual variety; in particular satisfying

1. $\forall xy(x \subseteq y \to D_w \subseteq D_v)$ (*monotonicity*)
2. $\forall xy(x \subseteq y \to \forall \vec{d} \in D_x(V_x(P, \vec{d}) = 1 \to V_y(P, \vec{d}) = 1)$ (*heredity*).

But other varieties, say with maps between the domains (cf. [Goldblatt, 1979]) would be suitable as well.

The 'de re/de dicto' interchange principles of Section 2.5 now have their obvious counterparts in the following quartetto:

1. $\neg \exists x Ax \to \forall x \neg Ax$,
2. $\forall x \neg Ax \to \neg \exists x Ax$,
3. $\exists x \neg Ax \to \neg \forall x Ax$,

4. $\neg\forall x Ax \to \exists x \neg Ax$.

The first three of these are universally valid on the present semantics. That they already hide quite some complexity is shown by the Gödel translation of (3):

$$\Box(\exists x \Box\neg\Box Ax \to \Box\neg\Box\forall x \Box Ax),$$

or

$$\Box(\exists x \Box\Diamond\neg Ax \to \Box\Diamond\exists x \Diamond\neg Ax).$$

No wonder that (3), e.g. does not define precisely the above monotonicity constraint on domains — even though its modal cousin $\exists x \Box Ax \to \Box\exists x Ax$ did.

The first really complex principle in Section 2.5 was the converse implication $\Box\exists x Ax \to \exists x \Box Ax$. We shall now investigate its intuitionistic cousin (4) — a rejected classical law.

EXAMPLE 89.

1. $\neg\forall x Ax \to \exists x \neg Ax$ implies that all domains are equal:

$$\forall xy(x \subseteq y \to D_x = D_y)$$

2. On frames with constant finite domain, $\neg\forall x Ax \to \exists x \neg Ax$ expresses the first-order condition that

$$\forall x\ (\exists! d\ d \in D_x \vee \forall y(x \subseteq y \to \forall z(x \subseteq z \to \exists u(y \subseteq u \wedge z \subseteq u)))).$$

Proof. Ad 1. Suppose that $x \subseteq y$, but $D_x \subsetneq D_y$. Make A true at y for all $d \in D_x$, and similarly at all $y' \supseteq y$. This stipulation defines an admissible assignment verifying $\neg\forall x Ax$ at x, while falsifying $\exists x \neg Ax$.

Ad 2. First, if $|D_x| = 1$, then trivially, $\neg\forall x Ax \to \exists x \neg Ax$ holds at x. (Recall that all domains are equal.)

Next, if $|D_x| > 1$, then one may argue as follows. If $\subseteq$ is *directed* above x in the above sense, then the assumption that $\exists x \neg Ax$ *fails* at x can be exploited to show that $\neg\forall x Ax$ must fail as well.

For, let $D_x = \{d_1, \ldots, d_k\}$. By the assumption, Ad_i will be true at some $x_i \supseteq x$ $(1 \leq i \leq k)$. Then, by successive applications of directedness, there will be found a common successor $y \supseteq x_1, \ldots, y \supseteq x_k$, where $\forall x Ax$ is true (by heredity). This falsifies $\neg\forall x Ax$ at x.

If on the other hand, for some $x, |D_x| > 1$ while $\subseteq$ is not directed above x, then, say, there exist $x_1 \supseteq x, x_2 \supseteq x$ without common successors. Then pick any object $d \in D_x$, making A true at x_1 and all its $\subseteq$- successors *for all objects except* d; while making A true at x_2 and all its $\subseteq$-successors *for* d *only.* This assignment verifies $\neg\forall x Ax$ at x, while falsifying $\exists x \neg Ax$. ■

Thus, a classical quantifier axiom may express an interesting purely relational constraint on $\subseteq$.

Now, intuitionists are fond of saying that (4) is valid for *finite* domains: as we have seen, however, it does impose constraints even then. They go on to say that an extrapolation to the infinite case would be illegitimate. At least, our principle becomes much more complex then.

THEOREM 90. *$\neg\forall xAx \to \exists x\neg Ax$ is not first-order definable in general.*

Proof. Consider the following structure, in which all worlds have a common domain N.

$$\begin{array}{ccccccccccc} & & & & & & \alpha & & & & \\ \bullet & \bullet & \bullet & \ldots & \longleftarrow & \bullet & \bullet & \bullet & \ldots & \longrightarrow & (\alpha < \omega_1) \\ 0 & 1 & 2 & & & -1 & 0 & +1 & & & \end{array}$$

i.e. $\langle W, \subseteq\rangle$ has the relational pattern of

$$\langle N \oplus (\omega_1 \odot \mathbb{Z}), \leq\rangle.$$

CLAIM. $\neg\forall xAx \to \exists x\neg Ax$ is true in this frame.

PROOF. Starting from any world x, assume that $\exists x\neg Ax$ fails. Then, for each $n \in N$, An must hold at some $(\alpha_n, k_n) > x$. As the cofinality of ω_1 exceeds ω, there exists some $\beta < \omega_1$ such that $(\beta, 0) > (\alpha_n, k_n)(n \in N)$. Now, by heredity, $\forall xAx$ must hold at $(\beta, 0)$ — whence $\neg\forall xAx$ is false at x. ■

Next, by the Löwenheim–Skolem theorem (as ever), this frame has *countable* elementary subframes. (Indeed, $\langle IN, \leq\rangle$ itself is one.) But in these, our principle may be falsified using some countable cofinal sequence $x_0, x_1, \ldots$ making $A0$ true from x_0 upward, $A1$ from x_1 upward, etcetera. As in earlier arguments, the conclusion of the theorem follows. ■

To finish this list of examples, it may be noted that a famous weaker variant of the above axiom does indeed define a first-order constraint.

EXAMPLE 91. Markov's Principle

$$\forall x(Ax \vee \neg Ax) \wedge \neg\neg\exists xAx \to \exists xAx$$

defines the relational condition

$$\forall x\exists y \supseteq x\ \forall z \supseteq y\ \forall d(Edz \to Edx).$$

Correspondence Theory remains surprising.

Post-Script: quantum logic.

Correspondences have not proved uniformly successful in intensional contexts. It seems only fair to finish with a more problematic example.

A possible worlds semantics for *quantum logic* was proposed in [Goldblatt, 1974]. Kripke frames are now regarded as sets of 'states' of some physical system, provided with a relation of 'orthogonality' ($\perp$). From its physical motivation, two pre-conditions follow for $\perp$, viz. *irreflexivity* and *symmetry.* But in addition, there is also a restriction to 'admissible ranges' for propositions, in the sense that these sets $X \subseteq W$ are to be *orthogonally closed*:

$$\forall x \in (W - X)\exists y \in (W - X)(\neg x \perp y \wedge \forall z \in X\ y \perp z).$$

The key truth clauses are those for conjunction (interpreted as usual), and negation, interpreted as follows:

$$\neg\varphi \text{ is true at } x \quad \text{if} \quad x \text{ is orthogonal to all } \varphi\text{-worlds}.$$

This semantics validates the usual principles for quantum logic, when $\vee$ is defined in terms of $\neg, \wedge$ by the De Morgan law. But, one key principle remains invalid, viz. the *ortho-modularity axiom*

$$p \leftrightarrow (p \wedge q) \vee (p \wedge \neg(p \wedge q)).$$

This axiom has a natural motivation in the Hilbert Space semantics for quantum logic — being the key stone in the representation of ortho-modular lattices as subspace algebras of suitable vector spaces. Thus, a minimal expectation would be that an enlightening correspondence is forthcoming with some constraint on the orthogonality relation $\perp$.

In reality, no such thing has happened. Quantum logicians pass onto *general* frames, into whose very definition validity of ortho-modularity has been built in. Despite this cover-up, the fact remains that the relational possible worlds perspective fails to do its correspondence duties here. A set-back, or an indication that facile *over-applicability* of Kripke semantics need not be feared for?

4 CONCLUSION

At a purely technical level, Correspondence Theory is an applied subject. Classical tools have been borrowed from model theory and universal algebra. In return to these mother disciplines, the subject offers a good range of (counter-)examples, as well as prospects for generalisability to other suitably chosen fragments of higher-order logic. (Cf. [van Benthem, 1983].)

From a more philosophical point of view, the whole enterprise may be described as finding out what possible worlds semantics really does for us.

It is one thing to make conceptual proposals, and another to really probe their depths. The systematic study of connections between intensional and classical perspectives upon possible world structures is an exploration of the benefits gained by the semantics. This chapter started with the observation that 'complex' modal axioms turned out to express 'simple' classical requirements (i.e. first-order ones). We have investigated the range and limits of this, and related phenomena. Especially these limits have become quite clear — and, with them, the limits of fruitful application of Kripke semantics. This philosophical conclusion holds for all semantics, of course. But we have earned the moral right to say it, through honest toil.

ACKNOWLEDGEMENTS

The classical introduction to a systematic modal model theory remains [Segerberg, 1971]. Some first applications of more sophisticated tools from classical model theory may be found in [Fine, 1975]. The algebraic connection was developed beyond the elementary level by L. Esakia, S. K. Thomason, R. I. Goldblatt and W. J. Blok. Two good surveys are [Blok, 1976] and [Goldblatt, 1979]. The proper perspective upon modal logic as a fragment of second-order logic was given in [Thomason, 1975]. An early appearance of correspondence theory proper is made in [Sahlqvist, 1975], full surveys are found in [van Benthem, 1983] for the case of modal logic and [van Benthem, 1985] for the case of tense logic. Other case studies are still in a preliminary state, with the exception of the intuitionistic treatise [Rodenburg, 1982].

University of Amsterdam

APPENDIX (1997)

This chapter first appeared in 1984. In the meantime, Modal Logic has evolved, but the basic structure of our original presentation remains valid. Therefore, we have left the old text unchanged, and merely added a short chronicle of further developments, including some answers to open questions. Generally speaking, correspondence methods have become a useful technical tool in pure and applied Modal Logic, without forming a major research area in their own right. A more principled motivation is given in van Benthem [1996a], where correspondence analysis is viewed as a central part in the philosophical quest for logical 'core theories' of semantic phenomena in language and computation. In particular, correspondences suggest the introduction of new many-sorted models, inducing decidable geometries of 'states' and 'paths' in the study of time and computation.

Extensions to Other Branches of Intensional Logic

The first significant extension of correspondence theory concerns *Intuitionistic Logic*. This involves the new feature that all valuations must be restricted to hereditary ones, leading only to formulas whose truth is preserved upward in the relational ordering. Rodenburg [1986] investigates this area in detail. In particular, he shows that the implication-conjunction fragment is totally first-order, whereas disjunctions can lead to non-first-orderness. Moreover, he introduces semantic tableau methods for explicit description of first-order correspondents. A final interesting feature is Rodenburg's analysis of intuitionistic Beth models which employ a second-order truth condition: a disjunction is true when its disjuncts 'bar' all future paths. These also turn out to be amenable to correspondence analysis, over two-sorted frames with both points and paths. Restricted valuations also occur with the ternary relational models of *Relevant Logic*. A full correspondence analysis is given in Kurtonina [1995], which analyses the special effects of working with features like distinguished points (actual worlds), non-standard connectives (including a new product conjunction), as well as the much poorer non-Boolean fragments found in categorial logics for grammatical analysis (cf. [van Benthem, 1991; Moortgat, 1996]). Further extensions have been made to *Epistemic Logic* [van der Hoek, 1992] and *Partial Logics* [Thijsse, 1992; Jaspars, 1994; Huertas, 1994]. Correspondence with restricted valuations for 'convex' propositions has also been proposed in standard *Temporal Logic* (cf. van Benthem [1983; 1986; 1995b]). But also, most axioms for richer interval-based versions have first-order 'Sahlqvist forms' [Venema, 1991]. Zanardo [1994] gives correspondences for modal-temporal models of branching space-time. Finally, correspondence methods have turned out very useful in *Algebraic Logic*. Venema [1991], Marx and Venema [1996] present a systematic study of relational algebra and cylindric algebra along these lines, pointing out the Sahlqvist form of most familiar algebraic axioms, and calculating their frame constraints on algebraic 'atom structures'. This establishes a much wider bridge between algebraic logic and modal logic than our earlier duality.

Restricted Frame Classes

Correspondence behaviour may change on special frame classes. In this chapter, we have looked at some effects of a restriction to transitive frames. But one can also investigate non-first-order frame classes. Van Benthem [1989a] considers *finite frames*, where, amongst others, the McKinsey axiom still defines a non-first-order condition. In this area, standard compactness-based model-theoretic techniques no longer work, and they must be replaced by a more careful combinatorial analysis with Ehrenfeucht-Fraïssé games of model comparison. (More generally, the finite model theory of modal logic is still undeveloped. Rosen [1995] proves some interesting transfer results, showing better finite model-theoretic behaviour than for first-order logic in general.) Doets [1987] takes up modal Ehrenfeucht games in great depth, investigating, amongst others, correspondence over *countable* and over *well-founded* frames. (For instance, the so-called Fine Axiom turns out to be first-order over countable frames.)

Complexity

This chapter contains some results on the (high) complexity of definability problems for monadic Φ_1^1–formulas. It turns out much harder to deal with the modal fragment of these. A lower bound for the complexity of first-orderness of modal formulas has been found in Chagrova [1991]: **M1** is *undecidable.* It seems likely that her methods (involving reductions of Minsky machine computation to correspondence statements) can also be made to yield non-arithmetical complexity. Conversely, undecidability of modal definability for first-order statements has been proved by Wolter [1993]: that is, **P1** is undecidable, too. A more general investigation of time and space complexity for modal logics, and the 'jumps' that may occur with different operator vocabularies, may be found in Spaan [1993]. It has improved decidability results for the so-called 'subframe logics' defined in Fine [1985], as well as 'transfer' of complexity bounds from components to compounds in poly-modal logics (cf. [Kracht and Wolter, 1991]).

Correspondence and Completeness

The main business of modal logic has been the search for completeness theorems over various frame classes. Correspondence theory bypasses this deductive information, focussing on direct semantic definability. Nevertheless, Kracht [1993] shows how the two enterprises can be merged, by a suitably generalized form of modal definability. Perhaps the most powerful result of this kind is the generalized Sahlqvist Theorem in Venema [1991], which shows that over suitably rich modal languages (possessing matched versions for each modality accessing all directions of its alternative relation), and allowing natural additional rules of inference beyond the minimal modal logic, the correspondence and the completeness version of the Sahlqvist Theorem converge in their proofs. The essential observation in the argument is as follows. In standard Henkin models for these richer systems, unlike in the standard case, all definable subsets employed in the correspondence proof (such as singletons or successor sets) are modally definable. Direct frame correspondences for modal rules of inference may be found in van Benthem [1985]. Over frames, the latter correspond to non-Φ_1^1 second-order formulas, but except for a few scattered observations in the literature, correspondence theory for modal rules of inference remains underexplored.

Duality with Algebraic Logic

Algebraic methods have been invaluable in finding key results on correspondence, such as the Goldblatt-Thomason characterization of the modally definable first- order formulas. Nevertheless, a purely model-theoretic reanalysis has been given in van Benthem [1993b], revolving around saturated models instead of descriptive frames. There is no definite preference here, as it is precisely the interplay between algebraic and model-theoretic viewpoints that remains fruitful. For new uses of correspondence methods in algebraic logic, as well as new set-theoretic representations for Boolean algebras with additional modal operators, see Marx [1995], Mikulas [1995]. For instance, Marx has an in-depth study of the duality between algebraic amalgamation and logical interpolation. The latter methods no longer employ simple binary relations as in the Jónsson-Tarski Stone representation, but more complex set-theoretic constructs. (Modal correspondences over finitary relations occur in van Benthem [1992], with a finite neighbourhood semantics for logic programs.) Developing a systematic correspondence theory over such generalized relational structures then becomes the next challenge.

Extended Modal Logics

Perhaps the most striking development in modal logic over the past ten years has been the systematic use of more powerful formalisms, with stronger modal operators over relational frames. A straightforward step is 'polymodal logic', which gives the same expressive power over frames with more alternative relations. Examples of the latter trend are the indexed modalities $< i >$ of propositional dynamic logic (cf. [Harel, 1984; Goldblatt, 1987; Harel *et al.*, 1998]), or n-ary modalities accessing $(n+1)$-ary alternative relations, as happens in relevant or categorial logics (cf. [Dunn, 2001; Kurtonina, 1995]). The correspondence theory of such extensions is straightforward, whereas there are interesting issues of 'transfer' for axiomatic completeness, finite model property, or computational complexity: cf. [Spaan, 1993; Fine and Schurz, 1996]. Transfer may depend very much on the connections between the various modalities. A case in point is modal predicate logic, whose theory has rapidly expanded over the past decade. Van Benthem [1993a] surveys some striking contributions by Ghilardi and Shehtman.

More interesting, from a correspondence point of view, is an increase in expressive power over the original binary relational frames. For temporal logic, the latter research line was initiated by Kamp's Theorem on functional completeness of the {Since, Until} language over continuous linear orders. In modal logic, the first systematic work emanated from the 'Sofia School': cf., e.g., [Gargov and Passy, 1990; Goranko, 1990], Vakarelov [1991; 1996]. These papers study addition of various new operators, such as a universal modality ranging over all worlds (relationally accessible or not), or various operations on poly-modalities, such as 'program intersection'. New frame constructions were invented to deal with these, such as 'duplication'. De Rijke [1992] investigates the 'difference modality' ("in at least one different world"), which has turned out to be useful and yet tractable. A more general program for extending modal logic (viewed as a general 'theory of information') occurs in van Benthem [1990] but the technical perspective is also clear in the pioneering paper Gabbay [1981]. Finally, de Rijke [1993] is an extensive model-theoretic investigation of definability and correspondence for extended modal languages, producing generalized versions for many results in this chapter (such as frame preservation theorems or effective correspondence algorithms). Still another angle on all this will follow below.

Alternatives: Direct Frame Theory

One may also analyze the frame content of modal logics more directly in terms of mathematical properties of graphs. Fine [1985] is a pioneer of this trend, emphasizing the good behaviour of 'subframe logics' which are complete for frame classes that are closed under taking subframes. (Such logics make no 'existential commitments'.) First-orderness is not a prominent consideration here: e.g., Löb's Axiom defines a simple subframe logic. Zakharyashev [1992; 1995] is a sophisticated study of modal logic from this viewpoint. Nevertheless, his direct classification of modal logics into three stages of frame preservation behaviour may again be reflected in second-order syntax and hence result in a form of correspondence theory at that higher level. A forthcoming monograph by Chagrov and Zakharyashev provides much more background, inluding references to earlier Russian sources (going back to Jankov in the sixties). Another excellent source, for many of the topics listed here, is the survey chapter [Chagrov *et al.*, 1996].

Models, Bisimulation and Invariance

Another noticeable shift of emphasis in the current literature leads away from frames to *models* as the primary objects of semantic interest. This move makes all of basic modal logic first-order, via our standard translation. The main questions then address what makes modal logics special as subspecies of first-order logic. In particular, what is the basic semantic invariance for basic modal logic, which should play a role like Ehrenfeucht games or 'partial isomorphism' in first-order model theory? A key result here is the semantic characterization of the modal fragment of first-order logic (modulo logical equivalence) as precisely those formulas in one free variable which are invariant for generated submodels and our 'zigzag relations' [van Benthem, 1976]. In modern jargon, this says that these formulas are precisely the ones *invariant for bisimulation*. The latter link was also developed in Hennessy & Milner [1985], which matches modal formalisms in different strengths with coarser or finer process equivalences. For up-to-date expositions of the resulting analogies between modal logics and computational process theories, cf. [van Benthem and Bergstra, 1995; van Benthem *et al.*, 1994], as well as various contributions in the volume [Ponse *et al.*, 1995]. This development has led to a new look at connections between modal formalisms and first-order logic. For instance, there are striking analogies between the meta-theories of both logics, whose precise extent and explanation is explored in de Rijke [1993], and Andréka, van Benthem & Németi [1998]. In particular, the latter paper investigates the hierarchy of *finite-variable fragments* for first-order logic as a candidate for a general account of modal logic (cf. [Gabbay, 1981; van Benthem, 1991] for this view). Typically, modal formulas need only two variables over worlds in their standard translation, temporal formulas only three, and so on. Finite-variable fragments are natural, and may be considered as functionally complete modal formalisms (cf. the insightful game-based analysis of Kamp's Theorem in Immerman & Kozen [1987]). Nevertheless, Andréka, van Benthem & Németi [1998] also turn up an array of negative properties, and eventually propose another classification for modal languages in terms of restricting atoms for *bounded quantifiers*. The resulting 'guarded fragments' can be analyzed much like the basic modal language, including analogous bisimulation techniques. In particular, these bisimulations now relate finite sequences of objects instead of single worlds, as in many-dimensional modal logics (cf. [Marx and Venema, 1996] for the theory of such formalisms). Their correspondence theory, taken with respect to natural generalized frame conditions for arbitrary first-order relations, still remains to be understood. [van Benthem, 1996b] is a general study of dynamic logics for computation and cognition, pursued via these techniques. One of its central concerns is expressive completeness of modal process logics vis-à-vis process equivalences like bisimulation.

Connections with Higher-Order Logic and Set Theory

From first-order correspondence, forays can be made into higher-order definability. Sometimes, this move is suggested by the modal language itself. E.g., in propositional dynamic logic, program iteration naturally translates into a countable disjunction of finite repetitions. Thus, translation into the *infinitary* standard language $L_{\omega_1\omega}$ seems the evident route. Infinitary frame correspondences were briefly considered in van Benthem [1983], and their modal model theory is explored in [de Rijke, 1993; van Benthem and Bergstra, 1995]. Of course, one may restore a balance here, and consider an infinitary modal counterpart of L_{ω}, allowing arbitrary set conjunctions and disjunctions, which would be the most natural formalism invariant for bisimulation. Barwise and Moss [1995] take this line, linking up truth on models and correspondence on frames. (Another perspective on infinitary modal logic is given in [Barwise and van Benthem, 1996].) Among a number of original results, they prove that a modal formula has all its infinitary substitution instances true in a model M iff it is true (in the usual second-order sense) on the frame collapse of that model taken with respect to the maximal bisimulation over M. As a direct consequence, frame correspondences for modal formulas imply model correspondences in infinitary modal logic. (The issue of good converses is still open). The original motivation for this type of investigation was that it relates modal logics to (non-well-founded) set theories. Linkages of this kind are further explored in d'Agostino [1995] which also raises the issue of more complex correspondences for modal axioms. For instance, she shows that the second-order Löb Axiom holds in a frame iff that frame is transitive while its collapse with respect to the maximal bisimulation is irreflexive. More generally, then, the interesting point about many correspondences is not that they must always reduce modal axioms to first-order ones, but rather the fact that they reformulate modal principles to any more perspicuous classical formalism. Another natural candidate of the latter kind is second-order monadic Φ^1_1 logic (cf. [Doets and van Benthem, 2001]). In particular, Doets [1989] shows how modal completeness theorems can sometimes be extended to cover this whole language. Moreover, many effective translation methods (see below) turn out to work for this broader language anyway. Finally, van Benthem [1989b] points out how first-order correspondence theory, suitably restated for second-order Φ^1_1 formulas, is a natural generalization which handles so-called computable forms of Circumscription in the AI literature (which involves reasoning from a second-order 'predicate-minimal' closure for first-order axioms; cf. [Lifshitz, 1985]).

Translations

Correspondence has become a conspicuous theme in the computational literature on theorem proving with intensional logics. A number of algorithms have been proposed, some of them rediscoveries of the Substitution Method and its ilk (cf. [Simmons, 1994]) and even much older results in second-order logic [Doherty, Łukasiewicz and Szalas, 1994], others working with new 'functional' translations better geared towards complete standard Skolemization and Resolution (cf. Ohlbach [1991; 1993]). One interesting feature of some of these algorithms is that they also produce useful equivalents for second-order modal principles. For instance, the typically non-first-order McKinsey Axiom gets a natural equivalent quantifying over both individual worlds and Skolem functions witnessing its (non-Sahlqvist) antecedent. Finally, we mention the use of set-theoretic interpretations of the standard translation in d'Agostino, van Benthem, Montanari & Policriti [1995], which read the universal modality as describing a power set. This translation also works with an explicit axiom system for general frames plus one axiom stating that the relational successors of any point in a frame form a set. This shift in perspective reduces theorem proving in modal logics to deduction in weak computational set theories. Many of these translations can also be formulated so as to deal with extended modal formalisms or larger fragments of second-order logic.

Designing New Logics

Finally, correspondence techniques have been used in 'deconstructing' standard logics and designing new ones. For instance, one can interpret first-order predicate logic over possible worlds models ('labelled transition systems') with assignments replaced by abstract states connected by abstract relations R_x modelling variable shifts. Then, standard predicate-logical validities turn out to express interesting frame properties, constraining possible computations, e.g., by Church-Rosser confluence properties (which match the first-order axiom $\exists y \forall x \phi \rightarrow \forall x \exists y \phi$). Moreover, one may want to impose certain restrictions on admissible valuations, such as 'heredity constraints' for axioms $\mathrm{Py} \rightarrow \forall \mathrm{xPy}$ or $\mathrm{Py} \rightarrow [y/x]\mathrm{Px}$ (van Benthem [1997; 1996b] have details). These abstract models reflect certain dependencies between admissible object values that may exist for individual variables. This theme is investigated more explicitly in [Alechina and van Benthem, 1993; Alechina, 1995], which design new generalized quantifier logics over 'dependence models', first proposed by Michiel van Lambalgen — where again the force of possible axioms is measured at least initially in terms of (Sahlqvist) frame correspondences. Related modal approaches to first-order logic are found in [Venema, 1991; Marx, 1995].

ADDED IN PRINT (1999)

Handbooks appear according to their own rhythms. Two years have elapsed since the updates were written for this Appendix. Here are a few further items of interest. D'Agostino [1998] contains new material on definability in infinitary modal logics, a topic also pursued further by Barwise and Moss. Meyer Viol [1995] has examples of correspondence for intuitionistic predicate logic showing how intermediate axioms can be quite surprising in their content. Hollenberg [1998] is an extensive study of definability, invariance and safety in modal process languages. Gerbrandy [1998] has interesting theorems on modal definability and bisimulation invariance in a setting of non-well-founded set theory, with applications to dynamic logic of epistemic updates. Grädel [1999] is an excellent survey of progress made on the program of decidable guarded first-order languages extending modal logic, including also fixed-point operators. Van Benthem [1998] is an up-to-date survey of the definability/correspondence paradigm, and the corresponding 'tandem approach' to modal and classical logics. Finally, two modern texts on modal logic that take correspondence seriously are Blackburn, de Rijke and Venema [1999] and van Benthem [1999].

BIBLIOGRAPHY

[d'Agostino *et al.*, 1995] G. d'Agostino, J. van Benthem, A. Montanari, and A. Policriti. Modal deduction in second-order logic and set theory. *Journal of Logic and Computation*, 7:251–265, 1997.

[d'Agostino, 1995] G. d'Agostino. Model and frame correspondence with bisimulation collapses, 1995. Manuscript, Institute for Logic, Language and Computation, University of Amsterdam.

[d'Agostino, 1998] G. d'Agostino. *Modal Logic and Non-well-founded Set Theory: Bisimulation, Translation and Interpolation.* PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 1998.

[Ajtai, 1979] M. Ajtai. Isomorphism and higher-order equivalence. *Ann Math Logic*, 16:181–233, 1979.

[Alechina and van Benthem, 1998] N. Alechina and J. van Benthem. Modal quantification over structured domains. In M. de Rijke, ed., *Advances in Intensional Logic*, pp. 1–28, Kluwer, Dordrecht.

[Alechina, 1995] N. Alechina. *Modal Quantifiers.* PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 1995.

[Andréka *et al.*, 1998] H. Andréka, J. van Benthem, and I. Németi. Back and forth between modal logic and classical logic. *Bulletin of the Interest group in Pure and Applied Logic*, 3:685–720, 1995. Revised version: Modal logics and bounded fragments of predicate logic. *Journal of Philosophical Logic*, 27:217–274, 1998.

[Barwise and Moss, 1995] J. Barwise and L. Moss. *Vicious Circles.* CSLI Publications, Stanford, 1995.

[Barwise and van Benthem, 1996] J. Barwise and J. van Benthem. Interpolation, preservation, and pebble games. Report ML-96-12, Institute for Logic, Language and Computation, University of Amsterdam, 1996. To appear in *Journal of Symbolic Logic*, 1999.

[Blackburn *et al.*, 1999] P. Blackburn, M. de Rijke and Y. Venema. *Modal Logic*, Kluwer, Dordrecht, 1999.

[Blok, 1976] W. J. Blok. *Varieties of interior algebras.* PhD thesis, Mathematical Institute, University of Amsterdam, 1976.

[Boolos, 1979] G. Boolos. *The Unprovability of consistency.* Cambridge University Press, Cambridge, 1979.

[Burgess, 1979] J. P. Burgess. Logic and time. *Journal of Symbolic Logic*, 44:566–582, 1979.

[Burgess, 1981] J. P. Burgess. Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic*, 22:76–84, 1981.

[Chagrova, 1991] L. Chagrova. An undecidable problem in correspondence theory. *Journal of Symbolic Logic*, 56:1261–1272, 1991.

[Chagrov *et al.*, 1996] A. Chagrov, F. Wolter and M. Zakharyashev. Advanced modal logic. School of Information Science, JAIST, Japan, 1996. To appear in *this Handbook.*

[Chang and Keisler, 1973] C. C. Chang and H. J. Keisler. *Model Theory.* North-Holland, Amsterdam, 1973.

[de Rijke, 1992] M. de Rijke. The modal logic of inequality. *Journal of Symbolic Logic*, 57:566–584, 1992.

[de Rijke, 1993] M. de Rijke. *Extending Modal Logics.* PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 1993.

[de Rijke, 1993a] M. de Rijke, ed. *Diamonds and Defaults*, Kluwer Academic Publishers, Dordrecht, 1993.

[de Rijke, 1997] M. de Rijke. *Advances in Intensional Logic*, Kluwer Academic Publishers, Dordrecht, 1997.

[Doets and van Benthem, 2001] K. Doets and J. van Benthem. Higher-order logic. In *Handbook of Philosophical Logic*, volume 1, 2nd edition, Kluwer Academic Publishers, 2001. (Volume I of 1st edition, 1983.)

[Doets, 1987] K. Doets. *Completeness and Definability. Applications of the Ehrenfeucht Game in Second-Order and Intensional Logic.* PhD thesis, Mathematical Institute, University of Amsterdam, 1987.

[Doets, 1989] K. Doets. Monadic Π_1^1 theories of Π_1^1 properties. *Notre Dame Journal of Formal Logic*, 30:224–240, 1989.

[Doherty,Lukasiewicz and Szalas, 1994] P. Doherty, W. Łukasiewicz and A. Szalas. Computing circumscription revisited: a reduction algorithm. Technical report LiTH-IDA-R-94-42, Institutioned for Datavetenskap, University of Linköping, 1994.

[Dunn, 2001] M. Dunn. Relevant logic. In *Handbook of Philosophical Logic*, volume 8, 2nd edition, Kluwer Academic Publishers, 2001. (Volume III of 1st edition, 1985.)

[Fine and Schurz, 1996] K. Fine and R. Schurz. Transfer theorems for multimodal logics. In J. Copeland, editor, *Logic and Reality. Essays on the Legacy of Arthur Prior.* Oxford University Press, Oxford, 1996.

[Fine, 1974] K. Fine. An incomplete logic containing **S4**. *Theoria*, 40:23–29, 1974.

[Fine, 1975] K. Fine. Some connections between elementary and modal logic. In S. Kanger, editor, *Proceedings of the 3rd Scandinavian Logic symposium.* North-Holland, Amsterdam, 1975.

[Fine, 1985] K. Fine. Logics containing K4. part II. *Journal of Symbolic Logic*, 50:619–651, 1985.

[Fitch, 1973] F. B. Fitch. A correlation between modal reduction principles and properties of relations. *Journal of Philosophical Logic*, **2**, 97–101, 1973.

[Gabbay, 1976] D. M. Gabbay. *Investigations in Modal and Tense Logics.* D. Reidel, Dordrecht, 1976.

[Gabbay, 1981] D. Gabbay. Expressive functional completeness in tense logic. In U. Mönnich, editor, *Aspects of Philosophical Logic*, pages 91–117. Reidel, Dordrecht, 1981.

[Gargov and Passy, 1990] G. Gargov and S. Passy. A note on Boolean modal logic. In P. Petkov, editor, *Mathematical Logic*, pages 311–321. Plenum Press, New York, 1990.

[Gerbrandy, 1998] J. Gerbrandy. *Bisimulations on Planet Kripke.* PhD theiss, Institute for Lgoic, Language and Computation, University of Amsterdam, 1998.

[Goldblatt and Thomason, 1974] R. I. Goldblatt and S. K. Thomason. Axiomatic classes in propositional model logic. In J. Crossley, editor, *Algebra and Logic.* Lecture Notes in Mathematics **450**, Springer, Berlin, 1974.

[Goldblatt, 1974] R. I. Goldblatt. Semantic analysis of orthologic. *Journal of Philosophical Logic*, 3:19–35, 1974.

[Goldblatt, 1979] R. I. Goldblatt. Metamathematics of modal logic. *Reports on Math. Logic*, 6:4–77 and 7:21–52, 1979.

[Goldblatt, 1987] R. Goldblatt. *Logics of Time and Computation.* Vol. 7 of CSLI Lecture Notes, Chicago University Press, 1987.

[Goranko, 1990] V. Goranko. Modal definability in enriched languages. *Notre Dame Journal of Formal Logic*, 31:81–105, 1990.

[Grädel, 1999] E. Grädel. Decision procedures for guarded logics. Mathematische Grundlagen der Informatik, RWTH Aachen, 1999.

[Grätzer, 1968] G. Grätzer. *Universal Algebra.* Van Nostrand, Princeton, 1968.

[Harel *et al.*, 1998] D. Harel, D. Kozen and J. Tiuryn. *Dynamic Logic.* 1998.

[Harel, 1984] D. Harel. Dynamic logic. In *Vol. II, Handbook of Philosophical Logic.* Reidel, Dordrecht, 1984.

[Henkin *et al.*, 1971] L. A. Henkin, D. Monk, and A. Tarski. *Cylindric Algebras I.* North-Holland, Amsterdam, 1971.

[Henkin, 1950] L. Henkin. Completeness in the theory of types. *Journal of Symbolic Logic*, 15:81–91, 1950.

[Hennessy and Milner, 1985] M. Hennessy and R. Milner. Algebraic laws for indeterminism and concurrency. *Journal of the ACM*, 32:137–162, 1985.

[Hollenberg, 1998] M. Hollenberg. *Logic and Bisimulation.* PhD Thesis, Institute of Philosphy, Utrecht University, 1998.

[Huertas, 1994] A. Huertas. *Modal Logics of Predicates and Partial and Heterogeneous Non-Classical Logic.* PhD thesis, Department of Logic, History and Philosophy of Science, Autonomous University of Barcelona, 1994.

[Humberstone, 1979] I. L. Humberstone. Interval semantics for tense logics. *Journal of Philosophical Logic*, 8:171–196, 1979.

[Immermann and Kozen, 1987] N. Immermann and D. Kozen. Definability with bounded number of bound variables. In *Proceedings 2nd IEEE Symposium on Logic in Computer Science*, pp. 236–244, 1987.

[Jaspars, 1994] J. Jaspars. *Calculi for Constructive Communication. The Dynamics of Partial States.* PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam and Institute for Language and Knowledge Technology, University of Tilburg, 1994.

[Jennings *et al.*, 1980] R. Jennings, D. Johnstone, and P. Schotch. Universal first-order definability in modal logic. *Zeit. Math. Logik*, 26:327–330, 1980.

[Kozen, 1979] D. Kozen. On the duality of dynamic algebras and Kripke models. Technical Report RC 7893, IBM, Thomas J. Watson Research Center, New York, 1979.

[Kracht and Wolter, 1991] M. Kracht and F. Wolter. Properties of independently axiomatizable bimodal logics. *Journal of Symbolic Logic*, 56:1469–1485, 1991.

[Kracht, 1993] M. Kracht. How completeness and correspondence theory got married. In M. de Rijke, ed. *Diamonds and Defaults*, pages 175–214. Kluwer Academic Publishers, Dordrecht, 1993.

[Kurtonina, 1995] N. Kurtonina. *Frames and Labels. A Modal Analysis of Categorial Inference.* PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam and Onderzoeksinstituut voor Taal en Spraak, Universiteit Utrecht, 1995.

[Lewis, 1973] D. Lewis. Counterfactuals and comparative possibility. *Journal of Philosophical Logic*, 2:4–18, 1973.

[Lifshitz, 1985] V. Lifshitz. Computing circumscription. In *Proceedings IJCAI-85*, pages 121–127, 1985.

[McTaggart, 1908] J. M. E. McTaggart. The unreality of time. *Mind*, **17**, 457–474, 1908.

[Marx and Venema, 1996] M. Marx and Y. Venema. *Multi-Dimensional Modal Logic.* Studies in Pure and Applied Logic. Kluwer Academic Publishers, Dordrecht, 1996.

[Marx, 1995] M. Marx. *Algebraic Relativization and Arrow Logic.* PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 1995.

[Meyer Viol, 1995] W. Meyer Viol. *Instantial Logic.* PhD thesis, Utrecht Institute for Linguistics OTS and Institute for Logic, Language and Computation, University of Amsterdam, 1995.

[Mikulas, 1995] S. Mikulas. *Taming Logics.* PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 1995.

[Moortgat, 1996] M. Moortgat. Type-logical grammar. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Language.* Elsevier Science Publishers, Amsterdam, 1996.

[Ohlbach, 1991] H-J Ohlbach. Semantics-based translation methods for modal logics. *Journal of Logic and Computation*, 1:691–746, 1991.

[Ohlbach, 1993] H. J. Ohlbach. Translation methods for non-classical logics. an overview. *Bulletin of the IGPL*, 1:69–89, 1993.

[Ponse *et al.*, 1995] A. Ponse, M. de Rijke, and Y. Venema, editors. *Modal Logic and Process Algebra—A Bisimulation Perspective.* Vol. 3 of CSLI Lecture Notes, Cambridge University Press, 1995.

[Rasiowa and Sikorski, 1970] H. Rasiowa and R. Sikorski. *The Mathematics of Metamathematics.* Polish Scientific Publishers, Warsaw, 1970.

[Rodenburg, 1982] P. Rodenburg. Intuitionistic correspondence theory. Technical report, Mathematical Institute, University of Amsterdam, 1982.

[Rodenburg, 1986] P. Rodenburg. *Intuitionistic Correspondence Theory.* PhD thesis, Mathematical Institute, University of Amsterdam, 1986.

[Rosen, 1995] E. Rosen. Modal logic over finite structures. Technical Report ML-95-08, Institute for Logic, Language and Computation, University of Amsterdam, 1995. To appear in *Journal of Logic, Language and Information*.

[Sahlqvist, 1975] H. Sahlqvist. Completeness and correspondence in the first- and second- order semantics for modal logic. In S. Kanger, editor, *Proceedings of the 3rd Scandinavian Logic Symposium*, pp. 110–143. North-Holland, Amsterdam, 1975.

[Segerberg, 1971] K. Segerberg. An essay in classical modal logic. *Filosofiska Studier*, 13, 1971.

[Simmons, 1994] H. Simmons. The monotonous elimination of predicate variables. *Journal of Logic and Computation*, 4:23–68, 1994.

[Smoryński, 1973] C. S. Smoryński. Applications of Kripke models. In A. S. Troelstra, editor, *Meta-mathematical Investigations of Intuitionistic Arithmetic and Analysis*, pages 329–391. Lecture Notes in Mathematics **344**, Springer, Berlin, 1973.

[Spaan, 1993] E. Spaan. *Complexity of Modal Logics*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 1993.

[Thijsse, 1992] E. Thijsse. *Partial Logic and Knowledge Representation*. PhD thesis, Institute for Language and Knowledge Technology, University of Tilburg, 1992.

[Thomason, 1974] S. K. Thomason. An incompleteness theorem in modal logic. *Theoria*, 40:30–34, 1974.

[Thomason, 1972] S. K. Thomason. Semantic analysis of tense logics. *Journal of Symbolic Logic*, 37:150–158, 1972.

[Thomason, 1975] S. K. Thomason. Reduction of second-order logic to modal logic I. *Zeit. Math. Logik*, 21:107–114, 1975.

[Vakarelov, 1991] D. Vakarelov. A modal logic for similarity relations in Pawlak information systems. *Fundamenta Informaticae*, 15:61–79, 1991.

[Vakarelov, 1996] D. Vakarelov. A modal theory of arrows. To appear in M. de Rijke, ed., 1996.

[van Benthem and Bergstra, 1995] J. van Benthem and J. Bergstra. Logic of transition systems. *Journal of Logic, Language and Information*, 3:247–283, 1995.

[van Benthem *et al.*, 1994] J. van Benthem, J. van Eyck, and V. Stebletsova. Modal logic, transition systems and processes. *Logic and Computation*, 4:811–855, 1994.

[van Benthem, 1976] J. F. A. K. van Benthem. *Modal correspondence theory*. PhD thesis, Instituut voor Grondslagenonderzoek, University of Amsterdam, 1976.

[van Benthem, 1978] J. F. A. K. van Benthem. Two simple incomplete modal logics. *Theoria*, 44:25–37, 1978.

[van Benthem, 1979a] J. F. A. K. van Benthem. Canonical modal logics and ultrafilter extensions. *Journal of Symbolic Logic*, 44:1–8, 1979.

[van Benthem, 1979b] J. F. A. K. van Benthem. Syntactic aspects of modal incompleteness theorems. *Theoria*, 45:67–81, 1979.

[van Benthem, 1980] J. F. A. K. van Benthem. Some kinds of modal completeness. *Studia Logica*, 39:125–141, 1980.

[van Benthem, 1981a] J. F. A. K. van Benthem. Intuitionistic definability, 1981. University of Groningen.

[van Benthem, 1981b] J. F. A. K. van Benthem. Possible worlds semantics for classical logic. Technical Report ZW-8018, Mathematical Institute, University of Groningen, 1981.

[van Benthem, 1983] J. van Benthem. *The Logic of Time*. Vol. 156 of *Synthese Library*, Reidel, Dordrecht, 1983. Revised edition with Kluwer Academic Publishers, Dordrecht, 1991.

[van Benthem, 1985] J. van Benthem. *Modal Logic and Classical Logic*. Vol. 3 of *Indices*, Bibliopolis, Napoli, and The Humanities Press, Atlantic Heights, NJ, 1985. Revised edition to appear with Oxford University Press.

[van Benthem, 1986] J. van Benthem. Tenses in real time. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 32:61–72, 1986.

[van Benthem, 1989a] J. van Benthem. Notes on modal definability. *Notre Dame Journal of Formal Logic*, 30:20–35, 1989.

[van Benthem, 1989b] J. van Benthem. Semantic parallels in natural language and computation. In H.-D. Ebbinghaus *et al.*, editor, *Logic Colloquium. Granada 1987*, pages 331–375. North-Holland, Amsterdam, 1989.

[van Benthem, 1990] J. van Benthem. Modal logic as a theory of information. In J. Copeland, ed., *Logic and Reality. Essays on the Legacy of Arthur Prior*, pp. 135–186. Oxford University Press, 1995.

[van Benthem, 1991] J. van Benthem. *Language in Action. Categories, Lambdas and Dynamic Logic.* Vol. 130 of *Studies in Logic*, North-Holland, Amsterdam, 1991.

[van Benthem, 1992] J. van Benthem. Logic as programming. *Fundamenta Informaticae*, 17:285–317, 1992.

[van Benthem, 1993a] J. van Benthem. Beyond accessibility: functional models for modal logic. In M. de Rijke, ed. *Diamonds and Defaults*, pp. 1–18. Kluwer, Dordrecht, 1993.

[van Benthem, 1993b] J. van Benthem. Modal frame classes revisited. *Fundamenta Informaticae*, 18:307–317, 1993.

[van Benthem, 1995b] J. van Benthem. Temporal logic. In D. Gabbay, C. Hogger, and J. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 4, pages 241–350. Oxford University Press, 1995.

[van Benthem, 1996a] J. van Benthem. Contents versus wrappings. In M. Marx, L. Pólos and M. Masuch, eds., *Arrow Logic and Multi-modal Logic*, pp. 203–219. CSLI Publications, Stanford, 1996.

[van Benthem, 1996b] J. van Benthem. *Exploring Logical Dynamics. Studies in Logic, Language and Information*, CSLI Pubications (Stanford) and Cambridge University Press, 1996.

[van Benthem, 1997] J. van Benthem. Modal foundations for predicate logic. Technical Report LP-95-07, Institute for Logic, Language and Computation, 1995. Appeared in *Bulletin of the IGPL*, 5:259–286, 1997. (R. de Queiroz, ed., Proceedings WoLLIC. Recife 1995) and E. Orlowska, ed., *Logic at Work. Memorial for Elena Rasiowa*, pp. 39–54, Studia Logica Library, Kluwer Academic Publishers, 1999.

[van Benthem, 1998] J. van Benthem. Modal loigc in two gestalts. To appear in M. de Rijke and M. Zakharyashev, eds. *Proceedings AiML-II, Uppsala*, Kluwer, Dordrecht, 1998.

[van Benthem, 1999] J. van Benthem. Intensional Logic. Electronic lecture notes, Stanford University, 1999. http://www.turing.wins.uva.nl/~johan/teaching/

[van der Hoek, 1992] W. van der Hoek. *Modalities for Reasoning about Knowledge and Quantities.* PhD thesis, Department of Computer Science, Free University, Amsterdam, 1992.

[Venema, 1991] Y. Venema. *Many-Dimensional Modal Logic.* PhD thesis, Institute for Logic, Language and Information, University of Amsterdam, 1991.

[White, 1981] M. J. White. Modal-tense logic incompleteness and the Master Argument. University of Arizona, 1981.

[Winnie, 1977] J. A. Winnie. The causal theory of space-time. In J. S. Earman *et al.*, editor, *Foundations of Space-Time Theories*, pages 134–205. University of Minnesota Press, 1977.

[Wolter, 1993] F. Wolter. *Lattices of Modal Logics.* PhD thesis, Zweites Mathematisches Institut, Freie Universität, Berlin, 1993.

[Zanardo, 1994] A. Zanardo. Branching-time logic with quantification over branches. The point of view of modal logic. Reprint 25, Institute for Mathematics, University of Padova, 1994.

[Zakharyashev, 1992] M. Zakharyashev. Canonical formuls for K4. Part I: Basic results. *Journal of Symbolic Logic*, 57:1377–1402, 1992.

[Zakharyashev, 1995] M. Zakharyashev. Canonical formulas for modal and superintuitionistic lgoics: a short outline. In M. de Rijke, ed. 1996.

Modal Quantification over Structured Domains

Johan van Benthem Natasha Alechina

1 Quantifiers as Modal Operators

1.1 Motivations

The semantics for quantifiers described in this paper can be viewed both as a new semantics for generalized quantifiers and as a new look at standard first-order quantification, bringing the latter closer to modal logic.

The standard semantics for generalized quantifiers interprets a monadic generalized quantifier Q as a set of subsets of a domain. For example, the quantifier "there are precisely two" is interpreted by the set of all subsets of the domain which contain precisely two elements. A formula $Qx\varphi$ is true in a model if the set of elements satisfying φ belongs to the interpretation of the quantifier; in our example, if there are precisely two elements satisfying φ. The existential quantifier can be treated as a generalized quantifier, too: it is interpreted as the set of all non-empty subsets of the domain. The universal quantifier is interpreted by the singleton set containing the whole domain.

The quantifiers listed so far are first-order definable in the following sense: they can be defined using ordinary quantifiers and equality. Many interesting generalized quantifiers are not first-order definable. The present study is motivated by the work of Michiel van Lambalgen (1991) on Gentzen-style proof theory for the quantifiers "for many" (its dual is interpreted as a non-principal filter), "for uncountably many" and "for almost all" (the latter contains all subsets of the domain which have Lebesgue measure 1). All those quantifiers are not first-order definable. They have Hilbert-style axiomatizations, but until lately no one believed that they can have a reasonable Gentzen-style proof theory. In order to devise such a proof theory, van Lambalgen used a translation of generalized quantifier formulas into a first-order language enriched with a predicate R of indefinite arity. $Qx\varphi(x, y_1, \ldots, y_n)$, where Q is a universal-type generalized quantifier (distributing over conjunction), is translated as $\forall x(R(x, y_1, \ldots, y_n) \to \varphi(x, y_1, \ldots, y_n))$, and its dual Q^d as $\exists x(R(x, y_1, \ldots, y_n) \wedge \varphi(x, y_1, \ldots, y_n))$ (all free variables displayed). Observe that this translation is reminiscent of the standard translation of modal formulas into first-order logic, with the sequence of free variables playing the role of the "actual world" and the quantifier ranging over the variables "accessible" from the given sequence. The idea behind such a translation is as follows. When generalized quantifiers are viewed as first-order operators (binding first-order variables), it becomes clear that a variable bound by a generalized quantifier cannot in general take any possible value. Its range is restricted, and this restriction can be defined using an accessibility relation. Then the elimination rule for Q with a premise $Qx\varphi(x, \bar{y})$ would introduce a variable $x_{\bar{y}}$ ranging over the set $\{x : R(x, \bar{y})\}$.

It turns out that some quantifier axioms correspond to first-order conditions on R in the following sense: any set of generalized quantifier formulas is consistent with the axiom if and only if the set of translations is consistent with the corresponding first-order condition on R. For example, $Qx\varphi \wedge Qx\psi \to Qx(\varphi \wedge \psi)$ corresponds in this sense to $R(x, \bar{y}\bar{z}) \to R(x, \bar{y})$.

In case such a correspondent exists, it is easier to find side conditions for elimination and introduction rules for the generalized quantifier satisfying the axiom. In the example above, the rule becomes

$$\frac{Qx\varphi(x,\bar{y})}{\varphi(x_{\bar{z}},\bar{y})}$$

where $\bar{y} \subseteq \bar{z}$ and $x_{\bar{z}}$ ranges over the set $\{x : R(x,\bar{z})\}$. Thus, we have a (generalized) modal logic for quantifiers here, which even exhibits some of the standard modal concerns (such as correspondence). Our aim in this paper is to investigate this general modal logic as such.

Another motivation for this enterprise is a modal-style modification of first-order quantification, aimed at obtaining a system which has some nice properties of modal logic not shared by first-order logic. In this sense, our approach relates to the one taken in Nemeti (1993) where non-standard models for first-order logic were introduced, yielding a decidable quantification theory by imposing restrictions on "accessible assignments". Let us rephrase our general idea. The Tarskian truth condition for the existential quantifier reads as follows:

$$M,[\bar{d}/\bar{y}] \models \exists x\varphi(x,\bar{y}) \;\Leftrightarrow\; \exists d \in D : M,[d/x,\bar{d}/\bar{y}] \models \varphi(x,\bar{y})$$

This may be viewed as a special case of a more general schema, when the element d is required in addition to stand in some relation R to $\bar{d}$ - where R is a finitary relation structuring the individual domain D:

$$M,[\bar{d}/\bar{y}] \models \Diamond_x\varphi(x,\bar{y}) \;\Leftrightarrow\; \exists d \in D :\; R(d,\bar{d}) \;\&\; M,[d/x,\bar{d}/\bar{y}] \models \varphi(x,\bar{y})$$

In the above-mentioned work on the generalized quantifiers "many", "uncountably many" and "almost all", R has properties which are common to different independence relations: linear independence in algebra, probabilistic independence, etc. But we can think of more general applications too, with domains being arranged in different levels of accessibility, or with procedures drawing objects in possible dependencies upon one another. One might read $R(d,\bar{e})$ as

- d can be constructed using $\bar{e}$,
- d is not "too far" from the e's,
- after you have picked up e's from the domain without replacing them, d is still available,

et cetera. Ordinary predicate logic then becomes the special case of flat individual domains admitting of "random access", whose R is the universal relation.

This semantics has some clear analogies with modal logic, with an existential generalized quantifier as an existential modality over some domain with, not a binary, but an arbitrary finitary "accessibility relation". As a consequence, we can apply standard ideas concerning modal completeness and correspondence to understand this broader concept of quantification.

Both motivations, generalized quantifiers and generalized first-order semantics, give rise to a variety of questions, including model theory, first-order completeness, canonicity, frame correspondence, definability of R-properties by quantifier axioms etc. In this paper we explore some of them.

1.2 Language and models

The language of the logic $EL(\exists, \Diamond)$ with a generalized quantifier is the ordinary language of first-order predicate logic with equality (without functional symbols) plus an existential generalized quantifier $\Diamond$. The notion of a w.f.f. is extended as follows: if φ is a w.f.f., then so is $\Diamond_x\varphi$. A universal dual of $\Diamond$ is defined as usual: $\Box_x\varphi =_{df} \neg\Diamond_x\neg\varphi$. We shall refer to the sublanguage without ordinary quantifiers as $EL(\Diamond)$.

$M = (D, R, V)$ is a *model* for $EL(\exists, \Diamond)$ if D and V are an ordinary domain and interpretation for first-order logic, and R is a binary relation between $d \in D$ and finite sequences $\bar{d}$ from D. Given the truth definition below, this is equivalent to considering a relation $R(d, D_0)$ between individual objects and finite *sets* D_0 of such objects.

The relation $M, v \models \varphi$ ("φ is *true* in M under assignment v") is defined as follows:

- $M, v \models P_i^n(x_{j1} \dots x_{jn}) \Leftrightarrow < v(x_{j1}) \dots v(x_{jn}) > \in V(P_i^n)$;
- $M, v \models \neg\varphi \Leftrightarrow M, v \not\models \varphi$;
- $M, v \models \varphi \wedge \psi \Leftrightarrow M, v \models \varphi$ and $M, v \models \psi$;
- $M, v \models \exists x\psi(x) \Leftrightarrow$ there exists a variable assignment v' which differs from v at most in its assignment of a value to x ($v' =_x v$) such that $M, v' \models \psi(x)$;
- $M, v \models \Diamond_x\psi(x, y_1, \dots, y_n) \Leftrightarrow$ there exists $v' =_x v$ such that $R(v'(x), v'(y_1), \dots, v'(y_n))$ and $M, v' \models \psi(x, \bar{y})$ where $\bar{y}$ are all (and just the) free variables of $\Diamond_x\psi$ listed in alphabetic order.

  It is easy to see that

- $M, v \models \Box_x\psi(x, \bar{y}) \Leftrightarrow$ if for all $v' =_x v$: $R(v'(x), v'(\bar{y})) \Rightarrow M, v' \models \psi(x, \bar{y})$.

We say that $M \models \varphi$ iff $M, v \models \varphi$ for all variable assignments v.

Let us define a *frame* (analogously to modal logic) $F = (D, R)$ as the underlying structure of a set of models with all possible interpretations of predicate letters. $F, v \models \varphi$ if $M, v \models \varphi$ for all models M on F. The formula φ is (globally) valid in F if, for all v, $F, v \models \varphi$ ("$F \models \varphi$").

This system resembles first-order logic in many respects, but no standard property can be taken for granted any more:

Monotonicity is restricted. Let for all variable assignments v $M, v \models \varphi(x_1, x_2) \rightarrow \psi(x_1, x_3)$ and for some assignment v $M, v \models \Diamond_{x_1}\varphi(x_1, x_2)$ (there exists $v' =_{x_1} v$ such that $R(v'(x_1), v(x_2))$ and $M, v' \models \varphi(x_1, x_2)$). It does not follow that $M, v \models \Diamond_{x_1}\psi(x_1, x_3)$, because although $M, v' \models \psi(x_1, x_3)$, it is not necessary that $R(v'(x_1), v(x_3))$ holds. Indeed, the general monotonicity rule

$$\frac{\Sigma \ \vdash \ \varphi(x, \bar{y}) \rightarrow \psi(x, \bar{z})}{\Sigma \vdash \Diamond_x\varphi(x, \bar{y}) \rightarrow \Diamond_x\psi(x, \bar{z}),}$$

with x not free in Σ, is invalid. We can accept only **Restricted Monotonicity**, where φ and ψ have the same free variables.

Extensionality is also restricted. Properties which hold for exactly the same objects, are no longer identical. Consider a property P which holds for a single object a: $\forall x(P(x) \equiv x = a)$. Let $R(a, \emptyset)$ and $\neg R(a, a)$. Then, $\Diamond_x P(x)$ is true and $\Diamond_x x = a$ is false.

Substitution therefore should also be restricted: only formulas with the same parameters can be substituted. We do not have in general that

$$D, R, V, v \models \varphi[\alpha/P] \ \Leftrightarrow\ D, R, V[P := [\alpha]_{M,v}], v \models \varphi.$$

2 Axiomatics and Completeness

We shall now develop the basic deductive calculus for our modal quantifier logic.

Definition 1 *The* minimal logic *for $EL(\exists, \Diamond)$ is a calculus of sequents $\Sigma \vdash \varphi$ satisfying the usual rules for first-order logic, including all Boolean principles, as well as the following quantifier rules:*

Restricted Monotonicity plus Distribution

$$\frac{\Sigma \ \vdash\ \varphi(x, \bar{y}) \to \bigvee_{i=1}^{i=n} \psi_i(x, \bar{y})}{\Sigma \vdash \Diamond_x \varphi(x, \bar{y}) \to \bigvee_{i=1}^{i=n} \Diamond_x \psi_i(x, \bar{y})}$$

where x is not free in Σ, and free variables are exactly those displayed (only x does not necessarily occur free in ψ_i). The convention here is that an empty disjunction is a falsum, both in the premise and the conclusion.

Alphabetic Variants

$$\vdash\ \Diamond_x \varphi(x, \bar{y}) \equiv \Diamond_z \varphi(z, \bar{y})$$

where z does not occur (free or bound) in $\varphi(x, \bar{y})$.

Here are some derivations in this system, corresponding to obvious validities given the above existential truth condition for the quantifier $\Diamond$:

1. $\vdash\ \bot \to \bot$

   $\vdash\ \Diamond_x \bot \to \bot$

   $\vdash\ \neg \Diamond_x \bot$

2. $\neg\varphi(\bar{y}) \vdash\ \varphi(\bar{y}) \to \bot$

   $\neg\varphi(\bar{y}) \vdash \Diamond_x \varphi(\bar{y}) \to \bot$

   $\vdash\ \Diamond_x \varphi(\bar{y}) \to \varphi(\bar{y})$, provided that x is not among the $\bar{y}$

3. Suppose that $\vdash \varphi \to \psi$ with x not free in ψ:

   Then:

   $\neg\psi \vdash\ \neg\varphi$

   $\neg\psi \vdash\ \varphi \to \bot$

   $\neg\psi \vdash\ \Diamond_x \varphi \to \bot$

   $\neg\psi \vdash\ \neg\Diamond_x \varphi$

whence $\vdash \Diamond_x\varphi \to \psi$.

4. An application of (3) is:

$$\vdash \ \Diamond_x\varphi \to \Diamond_x\varphi$$
$$\vdash \ \Diamond_x\Diamond_x\varphi \to \Diamond_x\varphi$$

5. Also,

$$\vdash \ \varphi \to \exists x\varphi$$
$$\vdash \ \Diamond_x\varphi \to \exists x\varphi$$

6. As a final illustration, we prove a useful principle for later reference, namely: $\vdash \neg\Diamond_z(\psi(z,\bar{y}) \wedge \neg\Diamond_x\psi(x,\bar{y}))$:

$$\vdash \psi(z,\bar{y}) \wedge \neg\Diamond_x\psi(x,\bar{y}) \to \psi(z,\bar{y})$$
$$\vdash \Diamond_z(\psi(z,\bar{y}) \wedge \neg\Diamond_x\psi(x,\bar{y})) \to \Diamond_z\psi(z,\bar{y})$$
$$\vdash \Diamond_z(\psi(z,\bar{y}) \wedge \neg\Diamond_x\psi(x,\bar{y})) \to \Diamond_x\psi(x,\bar{y})$$

and

$$\vdash \psi(z,\bar{y}) \wedge \neg\Diamond_x\psi(x,\bar{y}) \to \neg\Diamond_x\psi(x,\bar{y})$$
$$\vdash \Diamond_z(\psi(z,\bar{y}) \wedge \neg\Diamond_x\psi(x,\bar{y})) \to \neg\Diamond_x\psi(x,\bar{y})$$

(the latter step is as in example (3) above). Therefore,

$$\vdash \ \Diamond_z(\psi(z,\bar{y}) \wedge \neg\Diamond_x\psi(x,\bar{y})) \to \bot$$
$$\vdash \neg\Diamond_z(\psi(z,\bar{y}) \wedge \neg\Diamond_x\psi(x,\bar{y}))$$

Theorem 1 *The minimal logic is complete for universal validity.*

Proof. By a standard Henkin construction. The key point, as usual, is to create a maximally consistent set of formulas Σ - this time, adding suitable witnesses (new variables) for accepted formulas $\Diamond_x\varphi$:

If Σ_n is consistent with $\Diamond_x\varphi(x,\bar{y})$,

then add a *new* individual variable z with

1. $\varphi(z,\bar{y})$,
2. $\{\psi(z,\bar{y}) \to \Diamond_x\psi(x,\bar{y})|$ for all formulas $\psi\}$.

Claim. *This extension is consistent.*

Proof. Suppose it were inconsistent. Then, for some fresh variable z and some finite disjunction of formulas ψ_i:

$$\Sigma_n \vdash \ \varphi(z,\bar{y}) \to \bigvee_i(\psi_i(z,\bar{y}) \wedge \neg\Diamond_x\psi_i(x,\bar{y})).$$

Then also

$$\Sigma_n \vdash \Diamond_z \varphi(z, \bar{y}) \to \bigvee_i \Diamond_z(\psi_i(z, \bar{y}) \wedge \neg \Diamond_x \psi_i(x, \bar{y})).$$

Therefore, since $\vdash \Diamond_x \varphi(x, \bar{y}) \equiv \Diamond_z \varphi(z, \bar{y})$ (by Alphabetic Variants), $\{\Sigma_n, \ \Diamond_x \varphi(x, \bar{y})\}$ must be consistent with some

$$\Diamond_z(\psi_i(z, \bar{y}) \wedge \neg \Diamond_x \psi_i(x, \bar{y})).$$

But this contradicts the earlier derivability of the formula

$$\neg \Diamond_z(\psi(z, \bar{y}) \wedge \neg \Diamond_x \psi(x, \bar{y})).$$

□

Now construct the Henkin model as usual, and set

$$R(z, y_1, \ldots, y_n) \ \Leftrightarrow_{df} \ \forall \varphi : \varphi(z, \bar{y}) \in \Sigma \Rightarrow \Diamond_x \varphi(x, \bar{y}) \in \Sigma$$

(Note that we could have as well define $R(z, \{y_1, \ldots, y_n\})$ in the same way.) This definition may be compared with the usual introduction of the alternative relation R in completeness proofs for Modal Logic. To demonstrate the adequacy of the present Henkin model, all we have to prove is the following decomposition:

$$\Diamond_x \varphi(x, \bar{y}) \in \Sigma \ \text{ iff } \ \exists z : R(z, \bar{y}) \ \& \ \varphi(z, \bar{y}) \in \Sigma$$

From left to right, this is guaranteed by the above construction of Σ (through the addition of all formulas of the second kind). From right to left, this is a trivial consequence of the definition of R. □

If we look at the above completeness proof (and earlier examples of derivabilities), we see that no structural contraction rule or ordinary quantifier rules have been used. This observation (which is quite analogous with the situation in the minimal modal logic) motivates the conjecture the minimal logic without ordinary quantifiers is decidable. Indeed, the following theorem holds

Theorem 2 *The minimal logic without ordinary quantifiers and without equality is decidable.*

Proof. This is shown in Alechina (1994).

3 Model Theory

Now, to illustrate the semantical properties of modal quantifiers, we shall consider an analogue to the basic model-theoretic invariance relation of modal logic. In what follows, we talk about the language $EL(\Diamond)$ (without ordinary quantifiers).

Definition 2 *A* bisimulation $\mathcal{B}$ *between two models* $M_1 =< D_1, R_1, V_1 >$ *and* $M_2 = < D_2, R_2, V_2 >$ *is a family of partial isomorphisms* π *with the following properties:*

1 π *is a partial bijection with* $dom(\pi) \subseteq D_1$ *and* $ran(\pi) \subseteq D_2$*;*

2 If $\{d_1, \ldots, d_n\} \subseteq dom(\pi)$*, then for all predicate letters*

$$< d_1, \ldots, d_n > \in V_1(P^n) \ \Leftrightarrow < \pi(d_1), \ldots, \pi(d_n) > \in V_2(P^n)$$

($d_1, \ldots, d_n$ *are not necessarily distinct).*

3a If $D \subseteq dom(\pi)$ and $R_1(d, D)$, then there exists an element d' in D_2 such that $R_2(d', \pi[D])$ and $\{< d, d' >\} \cup \ \pi \in \mathcal{B}$.

3b If $D' \subseteq ran(\pi)$, $D' = \pi[D]$, and $R_2(d', D')$, then there exists an element d in D_1 such that $R_1(d, D)$ and $\{< d, d' >\} \cup \ \pi \in \mathcal{B}$. [1]

Invariance Lemma *If φ is a formula of $EL(\diamond)$ with the set of free variables $VAR(\varphi) \subseteq \{y_1, \ldots, y_n\}$, M_1 and M_2 are bisimilar models, and for all y_i $(1 \leq i \leq n)$ $v_1(y_i) \in dom(\pi)$ and $v_2(y_i) = \pi(v_1(y_i))$, $\pi \in \mathcal{B}$, then*

$$M_1, v_1 \models \varphi \ \Leftrightarrow \ M_2, v_2 \models \varphi$$

Proof. By induction on the length of φ.

- φ is a k-place predicate letter. - By clause (2) in the definition of bisimulation.
- $\varphi = (x = y)$. $M_1, v_1 \models x = y$ if and only if $v_1(x) = v_1(y)$. Since π is a function, and $v_1(x) \in dom(\pi)$, $\pi(v_1(x)) = \pi(v_1(y))$, that is, $v_2(x) = v_2(y)$ and $M_2, v_2 \models x = y$. Backwards: the same argument, using the fact that π is a bijection.
- $\varphi = \neg\psi$: by the inductive hypothesis,

  $$M_1, v_1 \models \psi \ \Leftrightarrow \ M_2, v_2 \models \psi$$

  and hence

  $$M_1, v_1 \models \neg\psi \ \Leftrightarrow \ M_2, v_2 \models \neg\psi.$$

- $\varphi = \psi_1 \wedge \psi_2$. Again, by the inductive hypothesis,

  $$M_1, v_1 \models \psi_1 \ \Leftrightarrow \ M_2, v_2 \models \psi_1$$

  $$M_1, v_1 \models \psi_2 \ \Leftrightarrow \ M_2, v_2 \models \psi_2$$

  and so,

  $$M_1, v_1 \models \psi_1 \wedge \psi_2 \ \Leftrightarrow \ M_2, v_2 \models \psi_1 \wedge \psi_2$$

- $\varphi = \diamond_x \psi(x, \bar{y})$. Assume $M_1, v_1 \models \diamond_x \psi(x, \bar{y})$. By the semantic truth definition, there exists an assignment v_1' which differs from v_1 at most in its assignment of value to x, such that $R(v_1'(x), v_1'(\bar{y}))$ and $M_1, v_1' \models \psi(x, \bar{y})$. By assumption, $y_1, \ldots, y_n \in dom(\pi)$. By clause 3a, there is $d' \in D_2$ with $R(d', \pi v_1'(\bar{y}))$, i.e. $R(d', v_2(\bar{y}))$ (since v_1' and v_1 agree on $\bar{y}$), and $\{< d, d' >\} \cup \ \pi \in \mathcal{B}$. Put $v_2' =_x v_2$, $v_2'(x) = d'$. Then, for the $\pi' \in \mathcal{B}$ which consists of π and the pair $< d, d' >$, $v_2'(x) = \pi'(v_1'(x))$, and for all y_i, $v_2'(y_i) = \pi(v_1'(y_i))$.

  By the inductive hypothesis, $M_2, v_2' \models \psi(x, \bar{y})$. But then $M_2, v_2 \models \diamond_x \psi(x, \bar{y})$. The same argument works backwards. $\square$

Continuing the analogy with modal logic, we define a translation of $EL(\diamond)$ formulas into the appropriate first-order logic, which is our original base language enriched with a dependence predicate R. The *standard translation ST* is defined as follows:

[1] Alternatively, we could restrict clause 3 to R-successors of the whole domain and range, while adding a further clause closing $\mathcal{B}$ under restrictions.

- $ST(P_i^n(t_1 \dots t_n)) := P_i^n(t_1 \dots t_n)$;
- $ST(t_1 = t_2) := (t_1 = t_2)$;
- ST commutes with classical connectives;
- $ST(\Diamond_x \varphi(x, \bar{y})) := \exists x(R(x, \bar{y}) \wedge ST(\varphi(x, \bar{y})))$.

Claim 1 *If φ is a formula of $EL(\Diamond)$, then*

$$M, v \models \varphi \;\Leftrightarrow\; M', v \models ST(\varphi),$$

for the classical model $M' = (D, V')$, where V' extends V to interpret the predicate R as R_M.

Definition 3 *The* modal formulas *(being those formulas which are standard translations of $EL(\Diamond)$ formulas) are the least set X of first-order formulas such that*

- *atomic formulas belong to X,*
- *if ψ_1 and ψ_2 are in X, then so are $\neg\psi_1$ and $\psi_1 \wedge \psi_2$,*
- *if $\varphi(x, \bar{y}) \in X$, then $\exists x(R(x, \bar{y}) \wedge \varphi(x, \bar{y}))$ is in X too.*

Theorem 3 *A first-order formula φ is equivalent to a modal formula if and only if it is preserved under bisimulation.*

Proof. The direction from left to right follows from Invariance Lemma above. For the converse, let φ be a first-order formula with variables $x_1, \dots, x_n$, preserved under bisimulation. We want to prove that it is equivalent to a modal formula.

Define the set $CONS_\Diamond(\varphi)$ as $\{\alpha : \alpha$ is a modal formula, $\varphi \models \alpha$ and the free variables of α are among $x_1, \dots, x_n\}$. If we can prove that

$$(*) \qquad CONS_\Diamond(\varphi) \models \varphi,$$

then we are done. For, by compactness, there will be some finite subset $\alpha_1, \dots, \alpha_m$ of $CONS_\Diamond(\varphi)$ with $\alpha_1, \dots, \alpha_m \models \varphi$. By definition, $\varphi \models \alpha_1, \dots, \alpha_m$. So, then φ is equivalent to $\alpha_1 \wedge \dots \wedge \alpha_n$, which is a conjunction of standard translations of $EL(\Diamond)$ formulas, i.e. a standard translation of the conjunction of those formulas.

Now we start proving $(*)$. Assume that for some model $M, v \models CONS_\Diamond(\varphi)$. We show that $M, v \models \varphi$. Let us denote the set of all modal formulas true in M and having free variables among $x_1, \dots, x_n$ as X_M. This is consistent with φ: for, if it is not, there is a finite set $\psi_1, \dots, \psi_k$ of formulas from X_M, such that $\bigwedge_i \psi_i \to \neg\varphi$. Then $\varphi \to \bigvee_i \neg\psi_i$. But $\bigvee_i \neg\psi_i$ is a modal formula (if every ψ_i is). Since it is a consequence of φ, it must be true in X_M. A contradiction.

Therefore there should be a model N for $\varphi \cup X_M$: say, $N, v' \models \varphi \cup X_M$.

Let $v(x_1) = d_1, \dots, v(x_n) = d_n$ in M and $v'(x_1) = d'_1, \dots, v'(x_n) = d'_n$ in N. Now, take ω-saturated elementary extensions $\mathcal{M}$ and $\mathcal{N}$ of M and N. We define a relation of bisimulation between $\mathcal{M}$ and $\mathcal{N}$ as follows:
(**) $\quad \mathcal{B}$ is the family of partial mappings π such that $\pi = \{(e_1, \pi(e_1)), \dots, (e_n, \pi(e_n))\}$ if

for all modal formulas ψ with at most free variables $x_1, \ldots, x_n$ and any two assignments v, v' with $v(x_i) = e_i$, $v'(x_i) = \pi(e_i)$ $(1 \leq i \leq n)$,

$$\mathcal{M}, v \models \psi \;\Leftrightarrow\; \mathcal{N}, v' \models \psi$$

To prove that $(**)$ indeed defines a bisimulation relation, we must check that the properties (1)–(3) hold for $\mathcal{B}$. Here, (1) is trivial. Case (2) is immediate, since atomic formulas are also standard translations of (atomic) formulas in $EL(\Diamond)$). Next, we check the zigzag clause 3a. Assume that $e_1, \ldots, e_k \in dom(\pi)$ and $R(e, e_1, \ldots, e_k)$. We must prove that there exists e' in $\mathcal{N}$ such that $R(e', \pi(e_1), \ldots, \pi(e_k))$ and $\{< e, e' >\} \cup \pi \in \mathcal{B}$. Take the set Ψ of all modal formulas with variables interpreted as $e, e_1, \ldots, e_k$ which are true in $\mathcal{M}$ under variable assignment v. We need an element e' in $\mathcal{N}$ such that all formulas in Ψ are true in $\mathcal{N}$ under v' when e' is assigned to the variable which was assigned e in $\mathcal{M}$. By saturation, it suffices to find such an e' for each finite subset Ψ_0 of Ψ. But these must exist, because the modal formula $ST(\Diamond_x \bigwedge \Psi_0(x, e_1, \ldots, e_k))$ holds in $\mathcal{M}$ and hence $ST(\Diamond_x \bigwedge \Psi_0(x, \pi(e_1), \ldots, \pi(e_k)))$ holds in $\mathcal{N}$. The appropriate check for the converse direction 3b is proved analogously.

Recall that $v(x_i) = d_i$ and $v'(x_i) = d'_i$, $1 \leq i \leq n$. We must also show that $\{< d_1, d'_1 >, \ldots, < d_n, d'_n >\} \in \mathcal{B}$. But this is so because for all modal formulas ψ with variables interpreted as $d_1, \ldots, d_n$ in M,

$$M, v \models \psi \;\Leftrightarrow\; N, v' \models \psi$$

(by the construction of N), and hence

$$\mathcal{M}, v \models \psi \;\Leftrightarrow\; \mathcal{N}, v' \models \psi.$$

Finally, since φ is invariant under bisimulation and $\{< d_1, d'_1 >, \ldots, < d_n, d'_n >\} \in \mathcal{B}$, $\mathcal{N} \models \varphi(d'_1, \ldots, d'_n)$ will now imply $\mathcal{M} \models \varphi(d_1, \ldots, d_n)$. Since $\mathcal{M}$ is an elementary extension of M, $M \models \varphi(d_1, \ldots, d_n)$, that is, $M, v \models \varphi(x_1, \ldots, x_n)$, and we are done. $\Box$

To conclude this section, we add some remarks about preservations properties for the full $EL(\exists, \Diamond)$ language. Since it includes the whole first-order language, bisimulation is obviously not enough to preserve all formulas:

Claim 2 *If φ does contain $\forall$ or $\exists$, bisimulation does not preserve truth.*

Proof. Let M_1 and M_2 be as follows:

$M_1 =< D_1, R_1, V_1 >$: $D_1 = \{d, d'\}$, $R_1 = \emptyset$, $V_1(P) = \{< d, d' >\}$;

$M_2 =< D_2, R_2, V_2 >$: $D_2 = \{e\}$, $R_2 = \emptyset$, $V_2(P) = \emptyset$.

Then $M_1, [d/x] \models \exists y P(x, y)$ and $M_2, [e/x] \not\models \exists y P(x, y)$. But at the same time, a bisimulation between these two models exists: $\mathcal{B} = \{< d, e >\}$. $\Box$

One can strengthen the above notion of bisimulation to preserve the full language, much as happens in modal logic extended with a "universal modality". The result is essentially the standard notion of "partial isomorphism" $\cong_p$ from abstract model theory.

4 Frame Correspondence

In this section and in the following one we extend the standard translation to $EL(\exists, \Diamond)$. An extra clause for ordinary quantifiers has to be added; as it is to be expected, ST commutes with ordinary quantifiers.

If a formula φ of $EL(\exists, \Diamond)$ is valid in a frame F (under an assignment v), then classically

$$F, v \models \forall P_i^n \dots \forall P_l^m ST(\varphi),$$

where $P_i^n, \dots, P_l^m$ are the predicate letters in φ. If this second-order formula has a first-order equivalent (containing only R and $=$), φ is called *first-order definable.* This means that if φ is true in all models over F, then R has the property defined by φ, and vice versa. Additional quantifier principles added to the minimal logic will now express special conditions on the relation R. One bunch of examples arises if we look at some properties of the standard existential quantifier $\exists$:

Unrestricted Distribution $\quad \Diamond_x(\varphi \vee \psi) \leftrightarrow \Diamond_x\varphi \vee \Diamond_x\psi.$

In one direction, this gives us unrestricted "Monotonicity" for $\Diamond$:

$$\Diamond_x\varphi \to \Diamond_x(\varphi \vee \psi).$$

This corresponds to the frame condition of

Upward Monotonicity $\quad R(x, \bar{y}) \to R(x, \bar{y}\bar{z})$

Proof. Suppose that $R(x, \bar{y})$. Define the following predicate:

$$P(u, \bar{v}) := \ u = x \ \wedge \ \bar{v} = \bar{y}.$$

We have $R(x, \bar{y}) \wedge P(x, \bar{y})$, whence $\Diamond_x P(x, \bar{y})$ holds. Therefore,

$$\Diamond_x(P(x, \bar{y}) \vee \bot(\bar{z}))$$

(where $\bot(\bar{z})$ is any contradiction involving $\bar{z}$): i.e., there exists d with $R(d, \bar{y}, \bar{z})$ and $P(d, \bar{y}) \vee \bot(\bar{z})$: the latter must be because $P(d, \bar{y})$: i.e. $d = x$, and hence $R(x, \bar{y}, \bar{z})$. $\quad \Box$

By a similar kind of argument, again making an appropriate substitution for the two predicates involved, the opposite direction

$$\Diamond_x(\varphi \vee \psi) \to \Diamond_x\varphi \vee \Diamond_x\psi$$

corresponds to the frame condition of

Downward Monotonicity $\quad R(x, \bar{y}\bar{z}) \to R(x, \bar{y})$

Together, these reduce the finitary relation R to an essentially *unary* "restriction" to the subdomain of all objects d satisfying the condition $R(d)$. It would also be of interest to see whether we can stop short of this, with quantifiers merely reducing the finitary relation R to a compound of *binary* ones (as happens in the generalized modal semantics for program operators proposed in van Benthem 1992).

Remark. Classical analogies may be slightly misleading here. E.g., the implication

$$\Diamond_x(\varphi \wedge \psi) \to \Diamond_x\varphi$$

(cf. $\Diamond(\varphi \wedge \psi) \to \Diamond\varphi$) expresses Downward Monotonicity, rather than the Upward Monotonicity of

$$\Diamond_x\varphi \to \Diamond_x(\varphi \vee \psi)$$

(cf. $\Diamond\varphi \to \Diamond(\varphi \vee \psi)$), even though the latter is equivalent with it in standard modal logic. Thus, it should in fact imply unlimited distribution - as may be seen using the available distribution in our minimal logic. In the latter calculus, "limited distribution" sanctions

1. $\Diamond_x(\varphi(x,\bar{y}) \vee \psi(x,\bar{z})) \to \Diamond_x((\varphi(x,\bar{y}) \wedge \top(\bar{z})) \vee (\psi(x,\bar{z}) \wedge \top(\bar{y})))$
2. $\Diamond_x((\varphi(x,\bar{y}) \wedge \top(\bar{z})) \vee (\psi(x,\bar{z}) \wedge \top(\bar{y}))) \to$
   $\to \Diamond_x(\varphi(x,\bar{y}) \wedge \top(\bar{z})) \vee \Diamond_x(\psi(x,\bar{z}) \wedge \top(\bar{y}))$,
3. from which the unlimited version $\Diamond_x(\varphi \vee \psi) \to \Diamond_x\varphi \vee \Diamond_x\psi$ follows by the above implication, passing to the appropriate conjuncts. $\Box$

Finally, the above unary relation gets trivialized to universality by the principle of

Instantiation $\quad \varphi \to \Diamond_x\varphi$

This corresponds to the frame condition $\forall x\forall y R(x,y)$ (provided that we assume non-empty individual domains, that is). The idea is this: let x,y be arbitrary, and let $P(x,y)$ hold of just these. We must have that $\Diamond_x P(x,y)$: i.e., some object d exists with $R(d,y)$ and $P(d,y)$: whence $R(x,y)$. $\Box$

Another source of examples is the analysis of various properties of the standard quantifier $\exists$ which are all lumped together as being "valid" in ordinary predicate logic, but which now become distinguishable as different properties of dependence. To be sure, such differences also become visible in other more sensitive semantics, such as those for intuitionistic predicate logic, or the logic of polyadic generalized quantifiers. Indeed, one concrete interpretation of the above structured domains would be the following:

individuals	pairs <world, individual>
dependence	$(w,x)R(v,y)$ iff $w \subseteq v$ & $y = x$,

inspired by standard possible worlds semantics for intuitionistic logic. We continue with one example of this kind:

Prenex operations $\quad \Diamond_x(\varphi \vee \psi) \leftrightarrow \varphi \vee \Diamond_x\psi$, where x not free in φ.

The direction $\to$ here turns out universally valid in case $\psi, \varphi \vee \psi$ have the same free variables, and hence derivable:

$\neg\varphi \vdash (\varphi \vee \psi) \to \psi$

$\neg\varphi \vdash \Diamond_x(\varphi \vee \psi) \to \Diamond_x\psi$

$\vdash \Diamond_x(\varphi \vee \psi) \to \Diamond_x\psi \vee \varphi$, Otherwise, it will enforce the earlier Downward Monotonicity: $R(x, \bar{y}\bar{z}) \to R(x, \bar{y})$. The direction $\leftarrow$ corresponds to the conjunction of $R(x, \bar{y}) \to R(x, \bar{y}\bar{z})$ and $\exists x R(x, \bar{y})$. $\square$

A comparison of quantifier axioms and similar modal axioms can also provide some interesting correspondences. For example, how would one write a quantifier version of the well-known K4-axiom: as

$$\Diamond_x \varphi \to \Diamond_x \Diamond_x \varphi$$

or with the more complex decoration

$$\Diamond_x \varphi \to \Diamond_y \Diamond_x \varphi?$$

The first one is universally valid, the second one defines

$$R(x, y\bar{z}) \to R(y, \bar{z})$$

(in case y is free in φ). Another direction is also possible: which quantifier principles correspond to well known properties of Kripke frames? Well-known examples are the three defining properties of equivalence relations:

Reflexivity $R(x, x)$

Transitivity $R(y, x) \wedge R(z, y) \to R(z, x)$

Symmetry $R(x, y) \to R(y, x)$

Fact. *The following principles are definable in* $EL(\exists, \Diamond)$*:*

- *Reflexivity corresponds to* $\Diamond_y x = y$*;*
- *Transitivity corresponds to* $\Diamond_y(\top(x) \wedge \Diamond_z(\top(y) \wedge P(z))) \to \Diamond_z(\top(x) \wedge P(z))$
- *Symmetry corresponds to* $\forall x \Box_y P(x, y) \to \forall y \Box_x P(x, y)$.

(Proofs will be given in Section 5 below.)

Some negative results concerning definability of first-order properties in $EL(\Diamond)$ alone can be obtained using frame constructions familiar from modal logic.

Definition 4 *Let* $F = < D, R >$ *be a frame and* $d_1, \ldots, d_n \in D$*. A subframe* $F' = < D', R' >$ *of* F *is* generated *by* $d_1, \ldots, d_n$ *if*

- D' *is the smallest subdomain of* D *containing* $d_1, \ldots, d_n$ *which is closed under accessibility, and*
- R' *is the restriction of* R *to* D'.

Theorem 4 *Let* F' *be a generated subframe of* F*,* v *a valuation restricted to the elements of* D'*, and* φ *a formula of* $EL(\Diamond)$*. Then*

$$F, v \models \varphi \;\Leftrightarrow\; F', v \models \varphi$$

(in other words, $EL(\Diamond)$*-formulas are invariant for generated subframes).*

Proof. For any pair of models $M =< F, V >$ and $M' =< F', V >$ the identity map from D' to D gives an obvious bisimulation, and we can apply our invariance results from Section 3. $\square$

Examples (Modal undefinability).

- $\exists x \neg R(x, x)$ is not definable by an $EL(\Diamond)$ formula. Consider

  $$F = < \{d_1, d_2\}, \{< d_1, d_1 >\} >,$$

  where it holds, and the generated subframe

  $$F' = < \{d_1\}, \{< d_1, d_1 >\} >,$$

  where it fails.

- $\forall x \forall y (x \neq y \;\Rightarrow R(x, y))$ is not definable in $EL(\Diamond)$. Consider the same two frames, but now in the opposite direction. $\square$

The language of $EL(\exists, \Diamond)$ with ordinary quantifiers added is much more powerful. For instance, $\exists x \neg R(x, x)$ is definable as $\exists x \neg \Diamond_y x = y$, and
$\forall x \forall y (x \neq y \Rightarrow R(x, y))$ as $\exists y (x \neq y \wedge P(x, y)) \to \Diamond_y P(x, y)$. Of course, a great deal of expressive power is due to the presence of identity in this language. Here is a more general result demonstrating this.

Theorem 5 *Every purely universal R-condition is* $\Diamond, \exists$*-definable.*

Proof. (Cf. Proposition 2.4 in de Rijke (1992a)). Consider any R-condition of the following form:

$$\forall y_1 \ldots \forall y_n \; BOOL(R, =, y_1, \ldots, y_n),$$

where "$BOOL$" is a purely Boolean condition. Introduce a predicate P_{y_i} for every universally quantified variable y_i, which holds exactly for y_i: $\exists ! x P_{y_i}(x)$. Define a translation * of first-order formulas with R into $EL(\exists, \Diamond)$, such that * commutes with Boolean connectives and $=$, where

$$(R(y, \bar{z}))^* = \; \Diamond_u (P_y(u) \wedge \top(\bar{z})).$$

Then the $EL(\exists, \Diamond)$ equivalent of the R-property will be

$$\exists ! x P_{y_1}(x) \wedge \ldots \wedge \exists ! x P_{y_n}(x) \to (BOOL(R, =, y_1, \ldots, y_n))^*$$

$\square$

Open problem *Are all first-order properties of* R $\exists, \Diamond$*-definable?*

We conjecture that the answer to this question is negative. A possible counterexample is the first-order formula $\exists x \forall y R(x, y)$.

There is a more general theory behind these various observations. The above axioms whose frame correspondences were analysed all had "Sahlqvist forms" in a suitably general sense, and the proof method depends on finding suitable "minimal substitutions". In the next section, we make this precise.

For now, we conclude with another aspect of modal frame correspondence. It is known that the question whether a modal axiom corresponds to a first-order condition on frames is undecidable (Chagrova (1991)). One would expect that the same holds for relational generalized quantifiers. And indeed we have this

Proposition. First-order correspondence for $EL(\Diamond)$ formulas is undecidable.

Proof. The idea is as follows. Let φ be a modal formula. It defines a first-order condition on frames if and only if $\forall P_1 \ldots \forall P_n ST(\varphi)$ has a first-order equivalent, where $P_1, \ldots, P_n$ are all predicate symbols in $ST(\varphi)$ and $ST(\varphi)$ is the standard translation of φ in the first-order language. Analogously for the generalized quantifier formulas. We find an effective fragment of the $EL(\Diamond)$ language whose first-order translations are effectively equivalent to the standard translations of modal formulas. Thus, a modal formula is first-order definable iff its $EL(\Diamond)$-counterpart is, and hence the correspondence problem for generalized quantifiers (in the latter language) is undecidable.

Consider the following translation $(\ )^i$ taking modal formulas to formulas of $EL(\Diamond)$ with one free variable w_i:

- $(p_n)^i = P_n(w_i)$
- commute with the Booleans
- $(\Box\varphi)^i = \Box_{w_{i+1}}(\top(w_i) \wedge (\varphi)^{i+1})$

The only thing to prove is that for every modal formula φ, $ST(\varphi)$ is provably equivalent in first-order logic to the standard translation of φ^0. (There is a minor difference which does not influence the result: modal R is the converse of our R.)

This may be shown by induction on complexity of φ. Let φ be a propositional variable. Then $ST(\varphi)$ is an atomic formula which is a standard translation of $(\varphi)^0$. In case φ is a negation or a conjunction, apply the inductive hypothesis (both standard translations commute with the Booleans). Let $\varphi = \Box\psi$. $ST(\Box\psi)[w] = \forall w'(R(w,w') \rightarrow ST(\psi)[w']) = \forall w'(R(w,w') \rightarrow \top(w) \wedge ST(\psi)[w']) =$ (by the inductive hypothesis) $= ST(\Box_{w'}(\top(w) \wedge ST(\psi)[w'])$. $\Box$

5 A Sahlqvist Theorem

Theorem 6 *All formulas of the "Sahlqvist form" $\bigwedge_i Qu_1 \ldots Qu_k(\varphi \rightarrow \psi)$, where Qu_j is either $\forall u_j$ or $\Box_{u_j}$, and*

1. *φ is constructed from*
   - *atomic formulas, possibly prefixed by $\Box_x$, $\forall$;*
   - *formulas in which predicate letters occur only negatively*

   using $\wedge, \vee, \Diamond_x, \exists$
2. *in ψ all predicate letters (except $=$) occur only positively*

are first-order definable.

Proof. If every conjunct is first-order definable, the whole conjunction is. Therefore without loss of generality we can concentrate on a formula of the form $Qu_1 \dots Qu_k(\varphi \to \psi)$. First we translate it into second-order logic:

$$\forall P_1^n \dots \forall P_l^m \forall u_1 \dots \forall u_k(\mathcal{R} \wedge ST(\varphi) \to ST(\psi)),$$

where $P_1^n \dots P_l^m$ are all the predicates in $\varphi \to \psi$ and $\mathcal{R}$ is a conjunction of R-statements corresponding to the $\Box$-quantifiers in the prefix. Then we remove all "empty" quantifiers (those binding variables not occurring in their scope), and rename bound individual variables in such a way that every quantifier gets its own variable which is distinct from any free variable occurring in the formula. Now it is possible to move all existential quantifiers occurring in positive subformulas of $ST(\varphi)$ to a prefix, using the following equivalences:

$$\exists x A(x) \vee \exists y B(y) \equiv \exists x \exists y (A(x) \vee B(y))$$

$$\exists x A(x) \wedge B \equiv \exists x (A \wedge B)$$

with the usual provisos on freedom and bondage. $ST(\varphi)$ has now been rewritten as

$$\exists y_1 \dots \exists y_m \varphi'.$$

Since ψ does not contain $y_1, \dots, y_m$ free, $\forall u_1 \dots \forall u_k(\mathcal{R} \wedge ST(\varphi) \to ST(\psi)$ is equivalent to

$$\forall x_1 \dots \forall x_n(\mathcal{R} \wedge \varphi' \to ST(\psi)),$$

where $x_1, \dots, x_n$ include $\bar{u}$ and $\bar{y}$.

Next, it would be convenient to get rid of the disjunctions in φ'. Let $\varphi' \equiv \phi_1 \vee \phi_2$.

$$\forall x_1 \dots \forall x_n((\mathcal{R} \wedge \phi_1) \vee (\mathcal{R} \wedge \phi_2) \to ST(\psi))$$

is equivalent to

$$\forall x_1 \dots \forall x_n(\mathcal{R} \wedge \phi_1 \to ST(\psi)) \wedge \forall x_1 \dots \forall x_n(\mathcal{R} \wedge \phi_2 \to ST(\psi)).$$

We can restrict attention to one of these conjuncts (if both components have a first-order equivalent, then so has their conjunction). So, assume that there are no disjunctions in the antecedent. Thus, we have a formula

$$\forall P_1^n \dots \forall P_l^m \forall x_1 \dots \forall x_n(\varphi' \to ST(\psi)),$$

where $P_1^n \dots P_l^m$ are all the predicates in $\varphi' \to ST(\psi)$, and φ' is a conjunction of "blocks" which are of one of the following forms:

1. standard translations of atomic formulas possibly preceded by universal and $\Box$-quantifiers,
2. R-statements,
3. formulas in which all predicate letters occur only negatively.

Next we rule out the use of negative formulas. The point is that $\varphi' \to ST(\psi)$ can always be rewritten as an implication whose antecedent does not contain negative formulas. Let $\varphi' = \phi_1 \wedge \phi_2$, where ϕ_2 is a negative formula. Then

$$\phi_1 \wedge \phi_2 \to ST(\psi)$$

is equivalent to

$$\phi_1 \to \neg\phi_2 \vee ST(\psi),$$

whose consequent contains only positive occurrences of predicate letters.

Let us denote the antecedent obtained (without negative formulas) φ^*. We shall now define the notion of *a minimal substitution* for every predicate letter in φ^*.

A predicate letter P_i^n can occur in φ^* more than once. Consider an occurrence $\bar{P}_i^n$ of P_i^n in φ^*. First we have to classify the variables of this occurrence (this is the only part where the present proof becomes different from the modal case). Let us assume that

- the variables which stand at the places $i_1, \ldots, i_m$ *in this occurrence* are existentially bound or free; let us denote them $x_1, \ldots, x_m$;

- the variables at the places $j_1, \ldots, j_k$ are universally bound by quantifiers which correspond to $\Box$-quantifiers in the original formula; let us call them $z_1, \ldots, z_k$;

- the rest of the variables is bound by ordinary universal quantifiers; let us call them $v_1, \ldots v_l$.

Before defining a minimal substitution we have to define the notion of an "R-condition" corresponding to the variable z_i:

1. Let $\Box_{z_1}$ be the first (leftmost) generalized quantifier in the sequence of quantifiers preceding $\bar{P}_i^n$, and before $\Box z_1$ the ordinary universal quantifiers $\forall v_1, \ldots, \forall v_s$ occur. Then the R-condition corresponding to z_1 will be $R(z_1, v_1, \ldots, v_s, \bar{x})$,

2. Let $\Box_{z_i}$ be the generalized quantifier following $\Box_{z_{i-1}}$ in our sequence (with some $\forall v_p, \ldots, \forall v_r$ possibly standing in between):

$$\ldots \Box_{z_{i-1}} \forall v_p \ldots \forall v_r \Box_{z_i} \ldots \bar{P}_i^n$$

   If the condition corresponding to z_{i-1} was $R(z_{i-1}, \bar{y})$, then the condition corresponding to z_i is $R(z_i, v_p, \ldots, v_r, z_{i-1}, \bar{y})$.

The minimal substitution $Sb(\bar{P}_i^n)$ for the occurrence of P_i^n in φ^* described above will be:

$P_i^n(u_1, \ldots u_n)$ is the conjunction of

1. $u_{i1} = x_1, \ldots, u_{im} = x_m$;

2. $\top(v_1), \ldots, \top(v_l)$;

3. $R(u_{\alpha_1}, \ldots, u_{\alpha_f})$, where $u_{\alpha_1}, \ldots, u_{\alpha_f}$ are the variables standing at the places $\alpha_1, \ldots, \alpha_f$, and in φ^* for these variables some R-condition (corresponding to one of the variables $z_1, \ldots, z_k$) hold.

Finally, we define

$$Sb(P_i^n, \varphi^*) = \bigvee Sb(\bar{P}_i^n)$$

for all occurrences of P_i^n in φ^*. [2]

[2]Note that we do not need existential quantifiers here to deal with iterations of $\Box$, as in modal logic; instead of $R^n(x, y)$, which is short for $\exists y_1(R(x, y_1) \wedge \ldots \wedge \exists y_{n-1} R(y_{n-1}, y))$, we have, for iterated modalities, $R(y_1, x) \wedge \ldots \wedge R(y, y_{n-1}, \ldots, y_1, x)$.

The result of substituting $Sb(P_i^n, \varphi^*)$ in $\forall x_1 \ldots \forall x_m(\varphi^* \to \psi')$, which we shall denote as

$$\forall x_1 \ldots \forall x_m(s(\varphi) \to s(\psi))$$

is our intended first-order equivalent, which contains no predicate symbols other than R and $=$. It is easy to see that it follows from the original Sahlqvist axiom, being an instantiation of a universal second-order formula

$$\forall P_i^n \ldots \forall P_l^m \forall x_1 \ldots \forall x_m(\varphi^* \to \psi').$$

We must prove the other direction to have an equivalence.

Assume that $\forall x_1 \ldots \forall x_m(s(\varphi) \to s(\psi))$ holds in some frame F under a variable assignment v. Assume, for some interpretation function V, that φ^* holds in $M =< F, V >$. To show that ψ' holds in the same model, we need the following two assertions:

Lemma 1 *For all M, v: $M, v \models \varphi^* \;\Rightarrow\; M, v \models s(\varphi)$*

Lemma 2 *Let $M, v \models \varphi^*$, and let $v(x_1) = d_1, \ldots, v(x_m) = d_m$. Define $V^*(P_i^n)$ as the set of all n-tuples which satisfy $Sb(P_i^n, \varphi^*)$ under v (that is, with $d_1, \ldots, d_m$ assigned to $x_1, \ldots, x_m$). Then*

$$V^*(P_i) \subseteq V(P_i).$$

From the first lemma it follows that $s(\varphi)$ also holds for V and v; and hence $s(\psi)$ holds. Since ψ' is positive, Lemma 2 (with the Monotonicity Lemma for classical logic) implies that $M, v \models \psi'$, as was to be shown.

Proof of lemma 1 φ^* has the form $\Psi \wedge \Gamma \wedge \Theta$, where Ψ is a conjunction of R-statements corresponding to the translations of $\Diamond$-quantifiers, Γ is a conjunction of atomic formulas, and Θ a conjunction of universally bound implications. It is easy to check that the two latter conjuncts turn into tautologies after substituting $Sb(P_i, \varphi^*)$ for every P_i in φ^*. It means that $\vdash\; s(\varphi) \equiv \Psi$, so it follows from any conjunction including Ψ.

Proof of lemma 2 (a.) Consider the case when the occurrence of P_i is in Γ. Every V which makes the formula true under v should include at least one tuple which satisfies the conditions from Ψ. Then it contains the tuple which satisfies $Sb(\bar{P_i})$. (b.) Let $\bar{P_i}$ be in Θ. Then it is of the form

$$\forall y_1 \ldots \forall y_{k+l}(\mathcal{R}_1 \wedge \ldots \wedge \mathcal{R}_k \to P_i(\bar{y}, \bar{x})),$$

where $\mathcal{R}_1 \ldots \mathcal{R}_k$ are the R-conditions corresponding to the generalized quantifiers. If φ^* is true under V and v, then this subformula is true, too, which means that $V(P_i)$ includes at least all tuples $< d_1, \ldots, d_n >$ for which the relation R holds between $\alpha_1, \ldots, \alpha_f$th members, for each of the k R-conditions. So, again it contains all tuples which satisfy $Sb(\bar{P}_i, \varphi^*)$. But if for every occurrence of P_i, the set of tuples satisfying $Sb(\bar{P}_i, \varphi^*)$ is a subset of $V(P_i)$, then also their union is in $V(P_i)$. Thus, $V^*(P_i) \subseteq V(P_i)$. $\Box$

Examples. Here is how the above Sahlqvist algorithm works on the earlier examples of reflexivity, transitivity and symmetry.

- *Reflexivity.* Consider $\Diamond_y x = y$. Its standard translation is

$$\exists y(R(y,x) \wedge x = y),$$

which is equivalent to $R(x,x)$.

- *Transitivity.* The standard translation of

$$\Diamond_y(\top(x) \wedge \Diamond_z(\top(y) \wedge P(z))) \to \Diamond_z(\top(x) \wedge P(z))$$

gives us

$$\forall P[\exists y(R(y,x) \wedge \top(x) \wedge \exists z(R(z,y) \wedge \top(y) \wedge P(z))) \to \exists u(R(u,x) \wedge \top(x) \wedge P(u))]$$

which can be rewritten in accordance with the Sahlqvist algorithm as

$$\forall P \forall y \forall z(R(y,x) \wedge R(z,y) \wedge P(z)) \to \exists u(R(u,x) \wedge P(u)))$$

The minimal substitution for $P(u)$ is $u = z$, so we obtain

$$\forall y \forall z(R(y,x) \wedge R(z,y) \wedge z = z \to \exists u(R(u,x) \wedge u = z),$$

which is a first-order equivalent of transitivity:

$$\forall y \forall z(R(y,x) \wedge R(z,y) \to R(z,x))$$

- *Symmetry.* The formula

$$\forall x \Box_y P(x,y) \to \forall y \Box_x P(x,y)$$

is translated as

$$\forall P(\forall x \forall y(R(y,x) \to P(x,y)) \to \forall y \forall x(R(x,y) \to P(x,y)))$$

The minimal substitution for $P(u,v)$ is $\top(u) \wedge R(v,u)$:

$$\forall x \forall y(R(y,x) \to \top(x) \wedge R(y,x)) \to \forall x \forall y(R(x,y) \to \top(x) \wedge R(y,x))$$

The antecedent becomes trivial:

$$\top \to \forall x \forall y(R(x,y) \to R(y,x))$$

which can again be written more elegantly as

$$\forall x \forall y(R(x,y) \to R(y,x)).$$

So far our illustations concerned finding first-order equivalents of modal formulas. Recall the reverse problem of defining first-order properties of R by modal formulas.

Fact. *All purely universal R-conditions can be defined using Sahlqvist formulas only.*

Proof. We show that the algorithm for defining R-properties in $EL(\exists, \Diamond)$ described in the Theorem 5 produces Sahlqvist formulas. First, $\bigwedge_i \exists! x P_{y_i}(x)$ is a Sahlqvist antecedent: it can be rewritten as

$$\bigwedge_i \exists y_i P_{y_i}(y_i) \wedge \bigwedge_i \forall x \forall z (P_{y_i}(x) \wedge P_{y_i}(z) \to x = z)$$

In the second conjunct, all predicate letters occur negatively (but when it is moved to the consequent in accordance with the Sahlqvist algorithm, those occurrences become positive).

Next, in the consequent we have $(BOOL(R, =, y_i))^*$, where some predicate letters again can occur negatively. Rewrite it as a conjunction of disjunctions of "atomic" statements $(\Diamond_u (P_{y_i}(u) \wedge \top(\bar{z})))$ and their negations:

$$\Phi \to \Psi_1 \wedge \ldots \wedge \Psi_n$$

The above expression is equivalent to the following conjunction:

$$(\Phi \to \Psi_1) \wedge \ldots \wedge (\Phi \to \Psi_n),$$

where each of Ψ_j's is a disjunction of atomic statements and their negations. Now move negations of atomic statements to the antecedents:

$$\Phi \to \neg \Diamond_u (P_y(u) \wedge \top(\bar{z})) \vee \Psi \quad \text{becomes} \quad \Phi \wedge \Diamond_u (P_y(u) \wedge \top(\bar{z})) \to \Psi$$

As a result, there are no negative occurrences of predicate letters in the consequents. □

Example (Symmetry revisited). Here is one more illustration of the preceding technique. Symmetry can be also defined "locally" using

$$P(x) \wedge \neg \exists x' (x' \neq x \wedge P(x')) \wedge Q(y) \wedge \neg \exists y' (y' \neq y \wedge Q(y')) \to$$

$$\to \neg \Diamond_u (P(u) \wedge \top(y)) \vee \Diamond_u (Q(u) \wedge \top(x)) :$$

The latter formula becomes

$$P(x) \wedge Q(y) \wedge \exists u (R(u, y) \wedge P(u)) \to \exists x' (x' \neq x \wedge P(x') \vee \exists y' (y' \neq y \wedge Q(y')) \vee$$

$$\vee \exists v (R(v, x) \wedge Q(v)),$$

or

$$\forall u [P(x) \wedge Q(y) \wedge R(u, y) \wedge P(u) \to \exists x' (x' \neq x \wedge P(x') \vee \exists y' (y' \neq y \wedge Q(y')) \vee$$

$$\vee \exists v (R(v, x) \wedge Q(v))].$$

The minimal substitutions are as follows:

$P(z) := \; z = x \vee z = u;$

$Q(z) := \; z = y.$

The resulting formula will be then

$$\forall u ((x = x \vee x = u) \wedge y = y \wedge R(u, y) \wedge (u = x \vee u = u) \to \exists x' (x' \neq x \wedge (x' = x \vee x' = u) \vee$$

$$\vee \exists y' (y' \neq y \wedge y' = y) \vee \exists v (R(v, x) \wedge v = y));$$

applying predicate logic gives

$$\forall u(R(u,y) \rightarrow \exists x'(x' \neq x \wedge x' = u) \vee R(y,x))$$

$$\forall u(R(u,y) \wedge \forall x'(x' = u \rightarrow x' = x) \rightarrow R(y,x),$$

which is equivalent to

$$\forall u(R(u,y) \rightarrow R(y,u))$$

□

6 Limitative Results

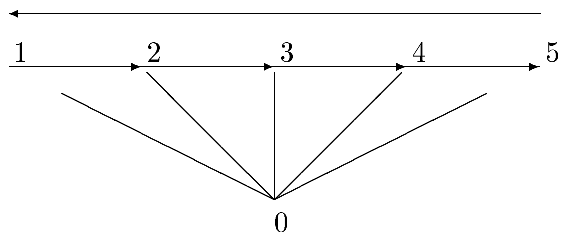
If a formula does not have the form described in our Sahlqvist Theorem, it may lack a first-order equivalent. The proof that a combination $\Box(\ldots \vee \ldots)$ in the antecedent can be fatal, is adapted from the analogous proof for modal logic (see van Benthem 1985, lemma 10.6).

Lemma 3 $\Box_x(\Box_y(P(y) \wedge \top(x,z)) \vee P(x)) \rightarrow \Diamond_x(\Diamond_y(P(y) \wedge \top(x,z)) \wedge P(x))$ *is not first-order definable.*

Proof. Define a class of frames F_n as follows:

- $D_n = \{0, 1, \ldots, 2n+1\}$;
- $R_n = \{< i, 0 >: 1 \leq i \leq 2n+1\} \cup \{< i+1, i, 0 >: 1 \leq i \leq 2n,\} \cup \{< 1, 2n+1, 0 >\}$.

Here is a picture illustrating this with $R(j,i,0)$ represented as "there is a line from 0 to i and an arrow from i to j":



For every n and V,

$$F_n, V, [z/0] \models \Box_x(\Box_y(P(y) \wedge \top(x)) \vee P(x)) \rightarrow \Diamond_x(\Diamond_y(P(y) \wedge \top(x)) \wedge P(x))$$

Indeed, the antecedent is true if

$$\forall x(R(x,z) \rightarrow \forall y(R(y,x,z) \rightarrow P(y)) \vee P(x));$$

that is, if for every i with $R(i,0)$ $P(i)$ is true or P holds for each j with $R(j,i,0)$. Each such i has exactly one "successor" j with $R(j,i,0)$ and "predecessor" k with $R(i,k,0)$. They form a chain which has by definition an odd number of members. That is why, if the antecedent is true, then P should hold for some pair of neighbours in this chain. But then the consequent is also true:

$$\exists x(R(x,z) \wedge \exists y(R(y,x,z) \wedge P(y)) \wedge P(x)).$$

Now, assume that our formula had a first-order equivalent. For arbitrary large n, it is consistent with the following set of first-order sentences describing the frames F_n:

$\forall x \forall y (R(x,y) \to \neg R(y,x))$

$\forall x \forall y \forall z (R(x,y,z) \to \neg R(y,x,z))$

$\exists ! z \forall y R(y,z)$

$\forall y (\exists ! x R(x,y,z) \wedge \exists ! u R(y,u,z))$

$\neg \exists x_1 \ldots \exists x_{2n} \exists y (R(x_2, x_1 y) \wedge \ldots R(x_{2n}, x_{2n-1}, y) \wedge R(x_1, x_{2n}, y))$.

The latter formula forbids "loops" of length less than $2n+1$; that is why it is true in F_k for all $k \geq n$.

By compactness, since each finite set of these formulas has a model for suitably large n, they also have a countable model simultaneously. But in all countable models with the above properties (which are isomorphic copies of $\mathbf{Z}$ with ternary R interpreted as $R(j,i,0) := S(j,i)$ and 0 being a fixed element preceding all other elements: $R(i,0)$ for all $i \neq 0$) the formula can easily be refuted by putting $P(i)$ iff $\neg P(i-1)$ and $\neg P(i+1)$. $\Box$

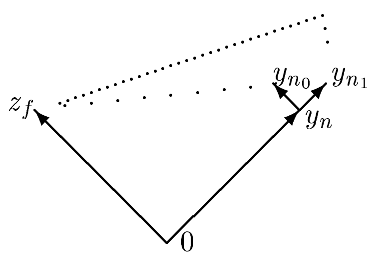
The same result holds for the combination $\Box_x \ldots \Diamond_y$ in the antecedent (the proof is analogous to the proof of lemma 10.2 in van Benthem 1985 for McKinsey axiom):

Lemma 4 $\Box_x \Diamond_y (P(y) \wedge \top(x,z)) \to \Diamond_x \Box_y (P(y) \wedge \top(x,z))$ *does not have a first-order equivalent.*

Proof. Consider the following class of models:

$D = \{0\} \cup \{y_n : n \in N\} \cup \{y_{n_i} : n \in N, i \in \{0,1\}\} \cup \{z_f : f : N \to \{0,1\}\}$;

$R = \{< y_n, 0 >: n \in N\} \cup \{< y_{n_i}, y_n, 0 >: n \in N, i \in \{0,1\}\} \cup \{< z_f, 0 >: f : N \to \{0,1\}\} \cup \{< y_{n_{f(n)}}, z_f, 0 >: n \in N, f : N \to \{0,1\}\}$



(Here an arrow from a to b describes $R(b,a)$, and the combination of arrows from a to b and from b to c - $R(c,b,a)$.)

Any model of this class validates the formula in question: assume

$$M, v = [z/0] \models \Box_x \Diamond_y (P(y) \wedge \top(x,z)).$$

This means that $\forall x (R(x,0) \to \exists y (R(y,x,0) \wedge P(y) \wedge \top(x,0))$ is true, which implies that $\forall n \exists i P(y_{n_i})$ holds. Since for every n either y_{n_0} or y_{n_1} satisfies P, we can choose f such that $P(y_{f(n)})$ for every n. Then the consequent is also true: $\exists x (R(x,0) \wedge \forall y (R(y,x,0) \to (P(y) \wedge \top(x,0)))$ (via $x = z_f$), whence

$$F, v = [z/0] \models \Box_x \Diamond_y (P(y) \wedge \top(x,z)) \to \Diamond_x \Box_y (P(y) \wedge \top(x,z))$$

M is obviously uncountable. Consider any countable elementary submodel M' of F which includes $0, y_n, y_{n_0}, y_{n_1}$ for all n. If our formula had a first-order equivalent, it would be true in M' . But it can be refuted there: since M' is countable, it does not contain some z_f. Put $y_{n_i} \in V(P)$ iff $i = f(n)$. Then the antecedent is still true (all elements which had a successor in P, still have it), but the consequent is false. □

Another limitation to the above result emerges when we try to obtain its natural generalization towards *completeness* of Sahlqvist logics. Here is a striking problem, due to Michiel van Lambalgen.

Example (Sahlqvist incompleteness).

Consider the following three axioms:

A1. $\Diamond_x x = x$;

A2. $\neg\Diamond_y x = y$;

A3. $\Diamond_x \varphi(x, \bar{y}) \to \Diamond_x(\varphi(x, \bar{y}) \vee \psi(x, \bar{z}))$

These properties are consistent (think of an interpretation for $\Diamond$ like "there exist at least two"). According to the Sahlqvist theorem, these axioms define the following properties of R:

R1. $\exists x R(x)$;

R2. $\neg R(x, x)$;

R3. $R(x, \bar{y}) \to R(x, \bar{y}, \bar{z})$;

But together R1–R3 imply $\perp$:

1. $R(x)$ - R1
2. $R(x) \to R(x, x)$ - R3
3. $R(x, x)$ - 1,2
4. $\neg R(x, x)$ R2
5. $\perp$

This example shows that the match between correspondence and completeness is not as good for modal quantifiers as it is for ordinary modal logic. A natural question arises, whether an analogue of the Sahlqvist's theorem can be proved for *correspondence for completeness*. The answer is given in Alechina and van Lambalgen (1994). In fact, the class of *weak Sahlqvist formulas* which have correspondents in the sense of completeness turns out to be a proper subclass of Sahlqvist formulas: namely, $\Diamond$ and $\exists$ quantifiers are not allowed in the antecedent.

7 Further Directions

In this paper we studied a number of properties of "modal" quantifiers, mostly their model theory and frame correspondence theory. Correspondence in the sense of completeness and its connection to the proof theory of generalized quantifiers is the main topic of Alechina and van Lambalgen (1994).

Another line of research is experimenting further with the truth definition so that to make the quantifier behave even more like a modal operator (and escape the incompleteness phenomenon described above). This can be achieved by replacing our accessibility relation between elements with one between assignments (van Benthem (1994)).

There are interesting connections between both lines of research sketched above and developments in algebraic logic (cf. Nemeti (1993)), namely in cylindric algebras. "Logically", cylindric algebras correspond to first order models with restricted sets of possible assignments. In van Benthem (1994) a set of conditions is found under which "abstract" frames for modal quantifiers (with the new truth definition) can be represented as "assignment frames", i.e. frames of such models. These and similar connections with algebraic logic form one of our directions of further research.

8 References

NATASHA ALECHINA & MICHIEL VAN LAMBALGEN (1994). Correspondence and completeness for generalized quantifiers. *Technical Note X-94-03,* ILLC, Univ. of Amsterdam.

NATASHA ALECHINA (1994). On one decidable generalized quantifier logic corresponding to a decidable fragment of first-order logic. *Submitted to the J. of Logic, Language and Information.*

JOHAN VAN BENTHEM (1985). *Modal logic and classical logic.* Bibliopolis, Napoli.

JOHAN VAN BENTHEM (1984). Modal correspondence theory. - D.Gabbay and F.Guenthner (eds.), *Handbook of Philosophical Logic,* **v.2**, Reidel, 167-247.

JOHAN VAN BENTHEM (1992). Logic as programming. - *Fundamenta Informaticae,* **17**(4), 285-317.

JOHAN VAN BENTHEM (1994). Modal state semantics. *Manuscript,* Univ. of Amsterdam.

LIDIA CHAGROVA (1991). An undecidable problem in correspondence theory. *Journal of Symbolic Logic,* **56**(4), 1261 - 1272.

MICHIEL VAN LAMBALGEN (1991). Natural deduction for generalized quantifiers. - J. van der Does & J. van Eijck (eds.), *Generalized Quantifier Theory and Applications,* Dutch Network for Language, Logic and Information, Amsterdam, 143-154.

ISTVAN NEMETI (1993). Decidability of weakened versions of first order logic. *Workshop Logic at Work, Amsterdam.*

MAARTEN DE RIJKE (1992a). The modal logic of inequality. - *Journal of Symbolic Logic,* **57**(2), 566-584.

MAARTEN DE RIJKE (1992b). How not to generalize Sahlqvist's theorem. - *Manuscript,* ILLC, Univ. of Amsterdam.

YDE VENEMA (1992). Many-dimensional modal logic. *PhD Thesis*, Univ. of Amsterdam.

Elementary Proof of the van Benthem-Rosen Characterisation Theorem

Martin Otto *

July 2003; revised May 04

Abstract

This note presents an elementary proof of the well-known characterisation theorem that associates propositional modal logic with the bisimulation invariant fragment of first-order logic. The classical version of this theorem is due to Johann van Benthem [2], its finite model theory analogue to Eric Rosen [8].

1 Introduction

The present proof of the van Benthem/Rosen characterisation theorem is uniformly applicable in both the classical and in the finite model theory scenario. While it is broadly based on Rosen's proof, it reduces the technical input from classical logic and the model theory of modal logics strictly to the use of Ehrenfeucht-Fraïssé games (for first-order, and for the modal variant). Furthermore the proof is constructive and the model constructions and accompanying analysis of games in the expressive completeness argument yield an optimal bound on the modal nesting depth in terms of the first-order quantifier rank. Despite this strengthening, the material becomes presentable in a highly self-contained manner, and can be covered even at the level of an introductory undergraduate course on logic and semantic games that covers the basic Ehrenfeucht-Fraïssé techniques.

Elsewhere this approach has been shown to extend and generalise to characterisations involving stricter forms of bisimulation (global and two-way, and to guarded bisimulation equivalence in transition systems) and corresponding extensions of basic modal logic in [6, 7]. A brief discussion is provided in Section 4.

*Mathematical Logic and Foundations of Computer Science, Department of Mathematics, Darmstadt University of Technology, otto@mathematik.tu-darmstadt.de

2 Notation & Preliminaries

2.1 Kripke structures and basic modal logic

2.1.1 Kripke structures

Consider Kripke structures or transition systems over a finite relational vocabulary consisting (w.l.o.g. for this note) of a single binary relation E and finitely many unary predicates $\mathbf{P} = (P_1, \ldots)$. We write $\mathcal{A} = (A, E^{\mathcal{A}}, \mathbf{P}^{\mathcal{A}})$ for a Kripke structure of this type, and typically indicate a distinguished element as in $\mathcal{A}, a$.

For $a \in A$, let $E^{\mathcal{A}}[a] = \{a' \in A \colon (a, a') \in E_j^{\mathcal{A}}\}$.

2.1.2 Basic modal logic

Denote as ML, or more specifically as $\mathrm{ML}[E; \mathbf{P}]$ propositional modal logic over this vocabulary. The formulae of $\mathrm{ML}[E; \mathbf{P}]$ are generated from $\top$, $\bot$ and the P in $\mathbf{P}$ allowing Boolean connectives and the modal quantifiers $\Box$ and $\Diamond$. The semantics is the usual one, with

$$\begin{array}{rlll} & \mathcal{A}, a \models \Box\varphi & \text{iff} & \mathcal{A}, a' \models \varphi \text{ for all } a' \in E^{\mathcal{A}}[a], \\ \text{and (dually)} & \mathcal{A}, a \models \Diamond\varphi & \text{iff} & \mathcal{A}, a' \models \varphi \text{ for some } a' \in E^{\mathcal{A}}[a]. \end{array}$$

We regard ML as a fragment of FO (or indeed FO^2, first-order logic with just two distinct variable symbols, x and y) via the standard translation based on

$$\begin{array}{rcl} [\Box\varphi]^*(x) & = & \forall y\big(Exy \rightarrow [\varphi]^*(y)\big), \\ [\Box\varphi]^*(y) & = & \forall x\big(Eyx \rightarrow [\varphi]^*(x)\big). \end{array}$$

We let ML_ℓ stand for the fragment of modal logic consisting of formulae whose nesting depth w.r.t. $\Box/\Diamond$ is at most ℓ. Note that the modal nesting depth coincides with the FO quantifier rank in terms of the standard translation.

2.1.3 Tree structures and their local relatives

A Kripke structure with distinguished element, $\mathcal{A}, a$, is called a *tree structure* if the underlying graph $(A, E^{\mathcal{A}})$ is a directed tree with root a in the graph theoretic sense: E is loop-free and every node is reachable from a on a unique E-path.

A tree structure is of *depth* ℓ if the lengths of paths is bounded by ℓ.

As an intermediary between arbitrary (finite) Kripke structures and tree structures, we consider (finite) structures that look like trees up to a certain depth from the distinguished node.

Generally, in a Kripke structure $\mathcal{A}$ let the *ℓ-neighbourhood* of $a \in A$ be the set $U^\ell(a)$ of all nodes reachable from a on (directed, forward) E-paths of length up to ℓ; $\mathcal{A}\!\restriction\! U^\ell(a)$ correspondingly denotes the substructure induced in restriction to $U^\ell(a)$.

We say that $\mathcal{A}, a$ is *ℓ-locally a tree structure* iff $\mathcal{A}\!\restriction\! U^\ell(a), a$ is a tree structure.

2.2 Bisimulation

Bisimulation equivalence between Kripke structures with distinguished nodes is denoted as in $\mathcal{A}, a \sim \mathcal{B}, b$. The corresponding approximations to level ℓ (as induced by the ℓ-round bisimulation game, see below) are denoted as in $\mathcal{A}, a \sim_\ell \mathcal{B}, b$.

The bisimulation game is played by players **I** and **II** over two Kripke structures $\mathcal{A}, a$ and $\mathcal{B}, b$, which each carry a pebble, initially placed on the distinguished elements a and b, respectively. In each round, the challenger, player **I**, moves the pebble in one of the structures forward along an E-edge, and player **II** has to respond by moving the other pebble along an E-edge in the opposite structure. It is player **II**'s task to maintain atomic equivalence throughout: **II** loses as soon as the currently pebbled nodes fail to agree on all monadic predicates (atomic propositions). Apart from that, players lose when they cannot move, for lack of E-edges. We say that **II** has a *winning strategy in the (infinite) bisimulation game* on $\mathcal{A}, a$ and $\mathcal{B}, b$, if she has a strategy to respond to any challenges from **I** without losing, indefinitely; **II** has a *winning strategy in the ℓ-round bisimulation game* on $\mathcal{A}, a$ and $\mathcal{B}, b$, if she has a strategy to respond to any challenges from **I** without losing for ℓ rounds. Then

- $\mathcal{A}, a$ and $\mathcal{B}, b$ are bisimilar, $\mathcal{A}, a \sim \mathcal{B}, b$, iff **II** has a winning strategy in the (infinite) bisimulation game on $\mathcal{A}, a$ and $\mathcal{B}, b$.
- $\mathcal{A}, a$ and $\mathcal{B}, b$ are ℓ-bisimilar, $\mathcal{A}, a \sim_\ell \mathcal{B}, b$, iff **II** has a winning strategy in the ℓ-round bisimulation game on $\mathcal{A}, a$ and $\mathcal{B}, b$.

The standard Ehrenfeucht-Fraïssé analysis of the bisimulation game yields the following.

Lemma 2.1. *Over the class of all Kripke structures of a fixed finite relational type:*

(i) *$\sim_\ell$ has finite index;*

(ii) *$\mathcal{A}, a \sim_\ell \mathcal{B}, b$ iff a in $\mathcal{A}$ and b in $\mathcal{B}$ are indistinguishable in ML_ℓ;*

(iii) *each $\sim_\ell$ equivalence class is definable by an ML_ℓ formula.*

A further few simple but useful properties of bisimulation equivalence are summarised in the following. In the first lemma we refer to the operation of *disjoint sums* or *disjoint unions* of relational structures: if $\mathcal{A}$ and $\mathcal{C}$ are structures of the same relational type, we denote as $\mathcal{A} + \mathcal{C}$ their disjoint sum (union), whose universe is the disjoint union of the universes and with all relations interpreted as in $\mathcal{A}$ and $\mathcal{C}$, respectively. The second lemma captures the local nature of $\sim_\ell$.

Lemma 2.2. *Bisimulation equivalence is insensitive to disjoint sums. If $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are of the same relational type, then $\mathcal{A}, a \sim \mathcal{B}, b$ iff $\mathcal{A} + \mathcal{C}, a \sim \mathcal{B}, b$.*

Lemma 2.3. (i) *$\mathcal{A}, a \sim_\ell \mathcal{B}, b$ iff $\mathcal{A} \restriction U^\ell(a), a \sim_\ell \mathcal{B} \restriction U^\ell(b), b$.*

(ii) *if $\mathcal{A}, a$ and $\mathcal{B}, b$ are both tree structures of depth ℓ, then ℓ-bisimulation coincides with bisimulation: $\mathcal{A}, a \sim_\ell \mathcal{B}, b$ iff $\mathcal{A}, a \sim \mathcal{B}, b$.*

The familiar and intuitive process of *unravelling* always guarantees bisimilar companions that are tree structures, albeit typically infinite ones. If only ℓ-local tree likeness is required, finite bisimilar companions are easily constructed for finite structures.

The *tree unravelling* $\mathcal{A}_a^*$ of $\mathcal{A}$ from a, is obtained as follows. The universe of $\mathcal{A}_a^*$ is the set of all (directed, forward) E-paths from a in $\mathcal{A}$. E is interpreted in $\mathcal{A}_a^*$ so that for each $m \in \mathbb{N}$, each path of length $m+1$ is an E-successor of its initial segment of length m. The unary predicates are interpreted in accordance with the projection $\pi \colon \mathcal{A}_a^* \to \mathcal{A}$ that maps each path to its last node.

Lemma 2.4. *Let $\mathcal{A}, a$ be a Kripke structure with distinguished node a.*

(i) *The tree unravelling of $\mathcal{A}$ from a, $\mathcal{A}_a^*$, is a tree structure that is bisimilar to $\mathcal{A}$ via the natural projection $\pi \colon \mathcal{A}_a^*, a \sim \mathcal{A}, a$.*

(ii) *For every $\ell \in \mathbb{N}$, the restriction of the tree unravelling $\mathcal{A}_a^*$ to depth ℓ, is a tree structure of depth ℓ that is ℓ-bisimilar to $\mathcal{A}, a$: $\pi \colon \mathcal{A}_a^* \restriction U^\ell(a), a \sim_\ell \mathcal{A}, a$.*

(iii) *For a finite Kripke structure $\mathcal{A}$ with distinguished node a, and $\ell \in \mathbb{N}$: there is a partial unravelling (to depth ℓ) that yields a finite bisimilar companion that is ℓ-locally a tree structure.*

Proof. (i) is obvious. For (ii) one may appeal to (i) of the previous lemma.

For (iii), take the tree unravelling $\mathcal{A}_a^*$ in restriction to $U^\ell(a)$, and identify each node b^* in $\mathcal{A}_a^* \restriction U^\ell(a)$ at distance ℓ from the root (a leaf in $\mathcal{A}_a^* \restriction U^\ell(a)$) with the node $b = \pi(b^*)$ in a fresh disjoint isomorphic copy of $\mathcal{A}$. $\square$

2.2.1 Bisimulation invariance

Definition 2.5. A formula $\varphi(x) \in \mathrm{FO}[E; \mathbf{P}]$ is *bisimulation invariant* iff, whenever $\mathcal{A}, a \sim \mathcal{B}, b$ then $\mathcal{A}, a \models \varphi$ iff $\mathcal{B}, b \models \varphi$.

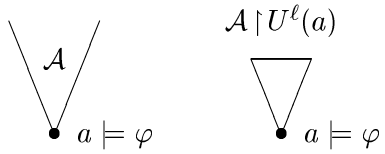
All formulae of ML are bisimulation invariant; in fact, $\varphi \in \mathrm{ML}_\ell$ is invariant under $\sim_\ell$. This is an immediate consequence of the modal Ehrenfeucht-Fraïssé analysis, or simply proved directly by syntactic induction on φ.

2.3 Locality

Gaifman's notion of *locality* [5] has been extensively studied in the first-order context, and in particular has proved to be a useful tool in finite model theory [4]. We here only need to make very limited use of the simple concept of ℓ-locality of a first-order formula in one free variable.

Definition 2.6. A property of Kripke structures with distinguished nodes $\mathcal{A}, a$—or a formula $\varphi(x)$ defining such a property—is *ℓ-local* iff whether or not it is satisfied in $\mathcal{A}, a$ only depends on $\mathcal{A} \restriction U^\ell(a), a$:

$$\mathcal{A}, a \models \varphi \quad \Leftrightarrow \quad \mathcal{A} \restriction U^\ell(a), a \models \varphi.$$



The following is a simple consequence of ℓ-bisimulation invariance, Lemma 2.1, and (i) in Lemma 2.3.

Observation 2.7. *Any* $\varphi \in \mathrm{ML}_\ell$ *is* ℓ*-local.*

3 The characterisation theorem

The goal of this note is a simple proof of the following characterisation theorem which goes through uniformly in the sense of finite model theory and classically. As an added benefit we obtain an optimal quantitative bound on quantifier ranks involved.

Theorem 3.1 (van Benthem/Rosen). *The following are equivalent for any* $\varphi(x) \in$ FO *of quantifier rank* q*:*

(i) $\varphi(x)$ *is invariant under bisimulation* [in finite Kripke structures].

(ii) $\varphi(x)$ *is logically equivalent* [over finite Kripke structures] *to a formula of* ML_ℓ*, where* $\ell = 2^q - 1$.

Note that the two readings—one classical, one finite model theoretic— really are two distinct theorems, a priori independent of each other. [1] Note that bisimulation invariance in finite structures does not imply bisimulation invariance over all structures: trivial examples are provided by formulae without finite models that happen not to be bisimulation invariant for infinite models.

Our proof proceeds in three stages. Note that even though we do not make this implicit in the statements, each statement is considered in its two readings: classically and in the sense of finite model theory.

Step 1 Any bisimulation invariant $\varphi(x) \in$ FO is ℓ-local for $\ell = 2^q - 1$ where $q = \mathrm{qr}(\varphi)$. This is proved with FO Ehrenfeucht-Fraïssé games, and as far as bisimulation is concerned rests on Lemma 2.2.

Step 2 Any bisimulation invariant $\varphi(x)$ that is ℓ-local, is even invariant under ℓ-bisimulation equivalence $\sim_\ell$. A simple bisimulation argument based on Lemmas 2.3 and 2.4 shows this.

Step 3 Any property invariant under ℓ-bisimulation equivalence is definable in ML_ℓ. This is a direct consequence of the Ehrenfeucht-Fraïssé analysis of bisimulation, based on Lemma 2.1 (iii).

[1] Two-variable first-order logic and two-pebble game equivalence illustrate this point. In that case, the classical characterisation theorem does not hold as a theorem of finite model theory, finite model property of FO^2 notwithstanding.

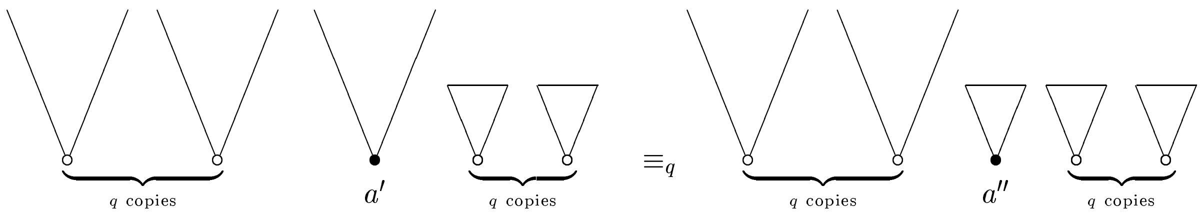
Proof sketch: step 1 Assume that $\varphi(x) \in \mathrm{FO}$ is bisimulation invariant, let $q = \mathrm{qr}(\varphi)$, and put $\ell := 2^q - 1$. To show that $\varphi(x)$ is ℓ-local, we consider any $\mathcal{A}, a$ and show that $\mathcal{A}, a \models \varphi$ iff $\mathcal{A} \restriction U^\ell(a), a \models \varphi$. As φ is bisimulation invariant, we may w.l.o.g. assume that $\mathcal{A} \restriction U^\ell(a), a \models \varphi$ is a tree of depth ℓ. [We may pass to a (finite partial) unravelling of $\mathcal{A}$, which is a bisimilar companion of $\mathcal{A}, a$ and whose restriction to $U^\ell(a)$ automatically is a bisimilar companion to $\mathcal{A} \restriction U^\ell(a), a \models \varphi$.] Here and in the core argument we use the fact that, due to bisimulation invariance, we may always replace a structure by some bisimilar companion without affecting (the property expressed by) φ.

If for some $\mathcal{A}', a' \sim \mathcal{A}, a$ and $\mathcal{A}'', a'' \sim \mathcal{A} \restriction U^\ell(a), a$, we can show $\mathcal{A}', a' \equiv_q \mathcal{A}'', a''$, we are done. For then

$$\begin{array}{lll} & \mathcal{A}, a \models \varphi & \\ \text{iff} & \mathcal{A}', a' \models \varphi & \text{(bisimulation invariance)} \\ \text{iff} & \mathcal{A}'', a'' \models \varphi & (\equiv_q \text{ equivalence}) \\ \text{iff} & \mathcal{A} \restriction U^\ell(a), a \models \varphi & \text{(bisimulation invariance)} \end{array}$$

For suitable $\mathcal{A}'$ and $\mathcal{A}''$, the equivalence $\mathcal{A}', a' \equiv_q \mathcal{A}'', a''$ can be established by exhibiting a strategy for player II in the q-round Ehrenfeucht-Fraïssé game.

As companions of $\mathcal{A}, a$ and $\mathcal{A} \restriction U^\ell(a), a$, respectively, we choose structures that are disjoint copies of sufficiently many isomorphic copies of $\mathcal{A}, a$ and $\mathcal{A} \restriction U^\ell(a), a$. Both structures involved will have q isomorphic copies of both $\mathcal{A}, a$ and $\mathcal{A} \restriction U^\ell(a), a$, and only distinguish themselves by the nature of the one extra component, in which live a' and a'', respectively. We indicate the two structures in the diagram below, with distinguished elements a' and a'' marked $\bullet$; the open cones stand for copies of $\mathcal{A}$, the closed cones for copies of $\mathcal{A} \restriction U^\ell(a)$. Clearly the structure on the left is bisimilar to $\mathcal{A}, a$, the one on the right bisimilar to $\mathcal{A} \restriction U^\ell(a), a$, by Lemma 2.2.



It suffices now to exhibit a strategy for player **II** in q rounds of the game on these structures. The game is started in the configuration with a single pebble in positions marked $\bullet$ in each of the two structures. The description of the strategy makes reference to a *critical distance* d_m, whose value for round m is

$$d_m = 2^{q-m}.$$

Starting with value $d_1 = 2^{q-1} = \lceil \ell/2 \rceil$ in the first round, this critical distance decreases by a factor $1/2$ in each round. Under the proposed strategy, **II** will always play *according* *to local context* if **I**'s move goes to an element within the critical distance from any already marked element; if **I**'s move is further than the critical distance away from all previously marked elements, we let **II** respond by marking *the same* element in one of the isomorphic copies of $\mathcal{A}$ or $\mathcal{A} \upharpoonright \ell$ that has not yet been touched by the game. (There are q many copies of each on each side, hence fresh ones are always available.)

The idea of *local context* works as follows. We think of the pebbles as belonging to disjoint clusters; initially we have just one single cluster consisting of the single elements with pebbles $\bullet$ on each side.

A pebble that is newly placed in round m joins an existing cluster if it is at most distance d_m away from one of the members of that cluster. Note that because of the shrinking d_m, no two clusters can ever be joined. Any elements of different clusters after round m are more than d_m apart.

Our strategy for **II** will have her maintain the condition that after completion of round m

> any two corresponding clusters are linked by an isomorphism that extends to all points within distance d_m of the members of the clusters.

If in round m, **I** places a new pebble further than d_m away from all previously marked elements, it forms a new cluster on its own, and **II**'s response into a new component makes sure that the same happens on the other side. If **I** places a new pebble to join one of the existing clusters, then **II** uses the isomorphism that comes with that cluster to respond with a matching element to join the corresponding cluster on the other side.

One checks that the above invariant is satisfied initially, and that the prescriptions for **II**'s moves are such that it is maintained in round m through all $m = 1 \dots, q$.

After q rounds, the local isomorphisms between clusters still guarantee that **II** wins.

Proof sketch: step 2 Let $\varphi(x)$ be ℓ-local and bisimulation invariant. Suppose $\mathcal{A}, a \sim_\ell \mathcal{B}, b$ and $\mathcal{A}, a \models \varphi$. We need to show that then also $\mathcal{B}, b \models \varphi$. Without loss of generality, we may assume that $\mathcal{A}$ and $\mathcal{B}$ are ℓ-locally tree structures. [If they are not, pass to (finite partial) unravellings, Lemma 2.4]

By ℓ-locality, $\mathcal{A}, a \models \varphi$ iff $\mathcal{A} \upharpoonright U^\ell(a), a \models \varphi$. Now $\mathcal{A}, a \sim_\ell \mathcal{B}, b$ iff $\mathcal{A} \upharpoonright U^\ell(a), a \sim_\ell \mathcal{B} \upharpoonright U^\ell(b), b$ iff (as both structures are now trees of depth ℓ) $\mathcal{A} \upharpoonright U^\ell(a), a \sim \mathcal{B} \upharpoonright U^\ell(a)$; Lemma 2.3.

Hence $\mathcal{A} \upharpoonright U^\ell(a), a \models \varphi$ iff $\mathcal{B} \upharpoonright U^\ell(b), b \models \varphi$. Therefore $\mathcal{B}, b \models \varphi$, by ℓ-locality again.

Proof sketch: step 3 If φ is invariant under $\sim_\ell$, we may use the ML_ℓ formulae that define $\sim_\ell$ equivalence classes, according to Lemma 2.1 (iii). Let $\chi_{\ell,\mathcal{A},a} \in \mathrm{ML}_\ell$ be the formula that defines the $\sim_\ell$ equivalence class of $\mathcal{A}, a$. Then φ is equivalent to the disjunction

$$\varphi \equiv \bigvee_{\mathcal{A},a \models \varphi} \chi_{\ell,\mathcal{A},a},$$

which is equivalent to a finite disjunction as $\sim_\ell$ has finite index.

This finishes the proof of the characterisation theorem.

Exercise 3.1. The exponential gap between the first-order quantifier rank q and modal quantifier rank $\ell = 2^q - 1$ cannot be avoided in general, as the example of formulae expressing that "there is an element satisfying p within distance $2^q - 1$" shows. Show that this is expressible in FO_q, but not by any formula in ML_m for $m < 2^q - 1$. [It can be expressed in modal quantifier rank $2^q - 1$.]

Observation 3.2. *There is an exponential succinctness gap between* FO *and* ML, *in expressing bisimulation invariant properties.*

4 Ramifications

In essence the technique outlined for basic modal logic above extends to other settings. We mention three distinct lines of variations and extensions.

Firstly, the classical statement of the theorem relativises to FO-definable classes of structures (as is clear also from the classical proof); but both the classical and the finite model theoretic versions also relativise to arbitrary bisimulation-closed classes.

Secondly, one can treat stronger variants of bisimulation equivalence, and in particular global forms of bisimulation, with a corresponding shift to technically more demanding locality arguments that need to apply uniformly across the entire structure rather than in a neighbourhood of the distinguished node.

Thirdly, one may want to consider other natural, more restricted classes of structures, rather than the class of all (or all finite) Kripke structures. Natural cases of interest include in particular classes of frames defined in terms of connectivity constraints, and in terms of classes of frames corresponding to classical modal theories.

4.1 Straightforward relativisations

Let $\mathcal{C}$ be a class of Kripke structures with distinguished elements. The notions of bisimulation invariance and of definability in ML give rise to corresponding notion in restriction to $\mathcal{C}$. For instance, $\varphi(x)$ is bisimulation invariant over $\mathcal{C}$ if for any two structures $\mathcal{A}, a$ and $\mathcal{B}, b$ from $\mathcal{C}$, $\mathcal{A}, a \sim \mathcal{B}, b$ implies that $\mathcal{A}, a \models \varphi$ iff $\mathcal{B}, b \models \varphi$. The classical proof of van Benthem's theorem uses compactness and saturation properties to establish indirectly that bisimulation invariance of $\varphi(x) \in \mathrm{FO}$ implies invariance ℓ-bisimulation for some ℓ. This argument clearly relativises to work within any class $\mathcal{C}$ defined by an FO theory.

The game oriented proof we gave above, on the other hand, is easily seen to work in restriction to (the finite structures within) any class $\mathcal{C}$ that is itself bisimulation closed.

Corollary 4.1. *Let $\mathcal{C}$ be closed under bisimulation, $\mathcal{C}_{fin}$ the class of finite structures within $\mathcal{C}$. Then $\varphi(x)$ is invariant under bisimulation over $\mathcal{C}$* [over $\mathcal{C}_{\mathrm{fin}}$] *iff $\varphi(x)$ is logically equivalent over $\mathcal{C}$* [over $\mathcal{C}_{\mathrm{fin}}$] *to a formula of* ML_ℓ, *where $\ell = 2^q - 1$.*

4.2 Stronger forms of bisimulation

Variations of this kind have been studied in [6, 7]. The following strengthenings of basic bisimulation equivalence are treated (described here in terms of the modifications in the corresponding bisimulation games):

(i) two-way bisimulation: **I** also has the option to move backward along E-edges, in which case **II** has to respond likewise.

(ii) global bisimulation: **I** can opt to move the pebble to a fresh start node anywhere in the structure, as can **II** in her response to such a move.

(iii) two-way and global bisimulation, $\approx$: both of the above.

It is entirely straightforward to adapt (the classical proof of) the classical characterisation theorem of van Benthem's to cover these variations. One naturally finds that these refined notions of bisimulation characterise within FO the following extensions of basic modal logic:

(i) two-way bisimulation: ML^{-}, ML with backward (past) modalities like $\Diamond^{-}\varphi(x) \equiv \exists y(Eyx \wedge \varphi(y))$.

(ii) global bisimulation: $\mathrm{ML}^{\forall}$, ML with a global modality, corresponding to unrestricted universal/existential quantification as in $\exists x\varphi(x)$ where $\varphi \in \mathrm{ML}$.

(iii) two-way and global bisimulation, $\approx$: $\mathrm{ML}^{-\forall}$, the combined extension by both of the above.

The global variants are technically interesting, because they require non-trivial locality arguments. We discuss the key case of $\approx$ invariance (global two-way bisimulation). On the one hand, the given FO formula $\varphi(x)$ is analysed in terms of Gaifman's locality theorem for FO, [5]. This allows us to determine locality parameters ℓ, m, q from φ such that whenever $\mathcal{A}, a$ and $\mathcal{B}, b$ agree

- on FO properties of quantifier rank q in the ℓ-neighbourhoods of a and b, respectively;
- on the quantifier rank q FO properties of systems of up to m many disjoint ℓ-neighbourhoods anywhere within $\mathcal{A}$ or $\mathcal{B}$, respectively;

then $\mathcal{A}, a \models \varphi$ iff $\mathcal{B}, b \models \varphi$.

In order to show the analogue of the crucial step 2, that then φ actually is invariant under $\approx_\ell$, one can construct, for arbitrary $\mathcal{A}, a \approx_\ell \mathcal{B}, b$, fully $\approx$ equivalent companion structures $\mathcal{A}^*, a \approx \mathcal{A}, a$ and $\mathcal{B}^*, b \approx \mathcal{B}, b$ that agree locally for ℓ, m, q in the above sense.

In the classical case, infinite companions $\mathcal{A}^*$ and $\mathcal{B}^*$ are admissible, and one can resort to bisimilar tree models obtained as suitable two-way unravellings, over which ℓ-two-way-bisimilarity then enforces local first-order equivalence.

In the finite model theory instance of the argument, one similarly seeks companion structures that are, at least ℓ-locally tree-like (acyclic), but at the same time need to be kept finite. This can be achieved with a construction of locally acyclic bisimilar coverings as developed in [6, 7]. These techniques have also been shown to extend to guarded bisimulation equivalence and to a characterisation of the guarded fragment GF of first-order logic, [1], over relational structures of width 2.

The general case for guarded bisimulation and the guarded fragment GF in relational structures with predicates of higher arities is the theme of ongoing investigations. In fact, it is currently open, whether the characterisation of GF as the guarded bisimulation invariant fragment of FO, due to Andréka, van Benthem and Németi [1], also obtains in the context of finite model theory. For this it would seem to be necessary to lift essential features of the construction of finite locally acyclic covers from the graph theoretic setting of bisimulations to the hypergraph theoretic setting of guarded bisimulations.

Further ramifications currently under investigation concern counting bisimulations (where the number of available successors of a certain kind matters) and modal logics with graded modalities.

4.3 Other natural classes of frames

These variations look at characterisation theorems for modal logics, of the above kind, over still more restricted classes of (finite) Kripke structures. Of particular interest from a transition system point of view are connected systems, as the existence of disconnected components (unreachable states) is often counterintuitive. Consider, for instance, the class of Kripke structures $\mathcal{A}, a$ that are *connected* in the sense that each node is reachable on some (forward, directed) E-path from a. In restriction to such connected frames, bisimulation equivalence coincides with global bisimulation equivalence, and correspondingly one expects a characterisation of the following kind.

Proposition 4.2. *The following are equivalent for every* $\varphi(x) \in \mathrm{FO}$*, both classically and in the sense of finite model theory:*

(i) φ *is invariant under bisimulation over the class of all* [finite] *connected Kripke structures.*

(ii) φ *is invariant under global bisimulation over the class of all* [finite] *connected Kripke structures.*

(iii) φ *is equivalent to a formula of* $\mathrm{ML}^{\forall}$ *over the class of all* [finite] *connected Kripke structures.*

Indeed, the constructive approach outlined above adapts to these settings to prove this proposition, as well as several other natural characterisation theorem of this kind. Interestingly, there seems to be no straightforward classical proof along the standard lines, as the underlying class of connected Kripke structures is not an elementary class, and compactness arguments are not directly available.

A number of other ramifications related to, for instance, frame conditions dealing with symmetry or transitivity requirements are being considered in ongoing joint work with A. Dawar [3]. Transitivity in particular is interesting from a technical point of view, as locality cannot be used in a straightforward manner.

References

1. H. Andréka, J. van Benthem, and I. Németi, *Modal languages and bounded fragments of predicate logic*, Journal of Philosophical Logic, 27 (1998), pp. 217–274.

2. J. van Benthem, *Modal Logic and Classical Logic*, Bibliopolis, Napoli, 1983.

3. A. Dawar and M. Otto, *Modal Characterisation Theorems over Connected Frames*, in preparation/unpublished note, 2003.

4. H.-D. Ebbinghaus and J. Flum, *Finite Model Theory*, Springer, 2nd ed., 1999.

5. H. Gaifman, *On local and nonlocal properties*, in Logic Colloquium '81, J. Stern, ed., North Holland, 1982, pp. 105–135.

6. M. Otto, *Modal and Guarded Characterisation Theorems over Finite Transition Systems*, Proceedings of 17th IEEE Symposium on Logic in Computer Science LICS '02, 2002, pp. 371–380.

7. M. Otto, *Modal and Guarded Characterisation Theorems over Finite Transition Systems*, Annals of Pure and Applied Logic, to appear, 2004.

8. E. Rosen, *Modal logic over finite structures*, Journal of Logic, Language and Information, 6 (1997), pp. 427–439.