Proofs and Expressiveness in Alethic Modal Logic

Maarten de Rijke   Heinrich Wansing

Version 3, February 2001

1 Introduction

Alethic modalities are the necessity, contingency,
possibility or impossibility of something being true.
Alethic means 'concerned with truth'.
[28, p. 132]

The above dictionary characterization of alethic modalities states the central notions of alethic modal logic: necessity, and other notions that are usually thought of as being definable in terms of necessity and Boolean negation: impossibility, contingency, and possibility. The syntax of modal propositional logic is inductively defined over a denumerable set of sentence letters p_0, p_1, p_2, ... as follows:

$$A ::= p \mid \neg A \mid (A \vee B) \mid \Box A$$

The other Boolean operations ($\top$, $\bot$, $\wedge$, $\supset$ and $\equiv$) are defined as usual. A formula $\Box A$ is read as "it is necessary that A", and:

"it is impossible that A"	is expressed as	$\Box\neg A$
"it is contingent that A"	is expressed as	$\neg\Box A$
"it is possible that A"	is expressed as	$\neg\Box\neg A$

Usually, "it is possible that A" is abbreviated $\Diamond A$, and if $\Diamond$ is primitive, $\Box A$ is abbreviated $\neg\Diamond\neg A$. Although one would expect that $\Box A$ implies A, the weakest system of normal modal propositional logic does not have $\Box A \supset A$ as a theorem. This is understandable from the point of view of the most prominent formal semantics for modal logic. The basic semantic intuition behind alethic modal logic is that $\Box A$ is true at a state ("possible world") s if and only if (iff) A is true at every state accessible from s. What exactly is meant by accessibility of t from s is deliberatively left open, to make room for various readings, like "t is compatible with the physical laws of s", "t is a conceptually possible alternative of s", "t lies in the future of s", or "t is an output-state of a terminating performance of some generic action in s". Clearly, if $\Box A$ is true at s iff A is true always in the future of s, the unprovability of $\Box A \supset A$ is intuitively correct.

Modal reasoning has been discussed by Aristotle already, and the idea of necessary truth as truth in all possible worlds is due to Leibniz, while its modern mathematical rendition goes back to Kripke. Over the past century modal logic has been used extensively to conceptualize and reason about a wide variety of modal and modal-like notions, some of which were mentioned above. To stay within the number of pages allotted to us, we have had to impose very drastic restrictions. First of all, our treatment is mainly logical or even mathematical. Second, we have decided to focus on two topics that, we think, are of relevance to anyone wanting to use modal logic for modeling and analyzing informal notions: *expressive power* (what can we say with the logic?) and *reasoning methods* (what are the implications of what we are saying?). In both cases we will focus on propositional modal logic; however, many interesting philosophical and mathematical phenomena and problems arise in modal predicate logic, and we will briefly touch on some of them.

More concretely, in Section 2, we survey the model theory of normal modal propositional logic and present basic notions and results of completeness and correspondence theory. Moreover, we indicate various ways of enhancing the expressive power of the language of alethic modal logic. In Section 3 we present an overview of two important types of proof systems for normal modal logics, namely labelled tableau systems and display calculi. Section 4 is concerned with several problems arising in modal predicate logic. We conclude this chapter with pointers to important survey articles and volumes on modal logic.

2 Model Theory

'Revolutionary' is an overused word, but no other word adequately describes the impact relational semantics (i.e., the concepts of frames, models, satisfaction, and validity that we are about to introduce) has had on the study of modal logic. Somewhere around 1960 modal logic was reborn as a new field, through the work of authors such as Hintikka, Kanger, and Kripke. Below we recall the basic concepts that came with these changes, and we discuss one of the key issues to which the new era gave rise: expressive power.

2.1 Basics

A *relational structure* is simply a tuple $(W, R_1, R_2, \ldots)$ consisting of a domain W and relations R_1, R_2, $\ldots$ on this domain. A *frame* for the propositional modal logic introduced in Section 1 is a relational structure $\mathcal{F} = (W, R)$ equipped with a single binary relation. A frame (W, R) is turned into a *model* $\mathcal{M} = (W, R, V)$ by equipping it with a *valuation* V, i.e., a function mapping proposition letters in the language to subsets of the domain W; note that models can be viewed as relational structures in a natural way, namely as structures of the form $(W,\ R,\ V(p_0),\ V(p_1),\ \ldots)$, consisting of a domain, a single binary relation R, and the unary relations given by V.

In spite of their mathematical kinship, frames and models are *used* very differently. Frames are essentially mathematical pictures of ontologies or structural properties that are more or less invariant across situations, while the unary relations provided by valuations decorate frames with contingent information.

Definition 1 Let w be a state in a model $\mathcal{M} = (W, R, V)$. We inductively define the notion of a formula A being *true* in $\mathcal{M}$ at w (notation: $\mathcal{M}, w \models A$) as follows:

$$
\begin{array}{rll}
\mathcal{M}, w \models p & \text{iff} & w \in V(p) \\
\mathcal{M}, w \models \neg A & \text{iff} & \text{not } \mathcal{M}, w \models A \\
\mathcal{M}, w \models A \vee B & \text{iff} & \mathcal{M}, w \models A \text{ or } \mathcal{M}, w \models B \\
\mathcal{M}, w \models \Box A & \text{iff} & \text{for all } v \in W \text{ with } wRv \text{ we have } \mathcal{M}, v \models A
\end{array}
$$

It follows from this definition that $\mathcal{M}, w \models \Diamond A$ iff for some $v \in W$ with wRv we have $\mathcal{M}, v \models A$. Note also that the notion of truth is *local*: formulas are evaluated at some particular state w. Moreover, $\Box$ and $\Diamond$ both work locally: only states R-accessible from the current one can be explored by our operators.

A formula A is *globally* or *universally true* in a model $\mathcal{M}$ (notation: $\mathcal{M} \models A$) if it is true at all states in $\mathcal{M}$.

Finally, these notions can also be lifted to sets of formulas Σ: $\mathcal{M}, w \models \Sigma$ if $\mathcal{M}, w \models A$ for every $A \in \Sigma$; and $\mathcal{M} \models \Sigma$ if $\mathcal{M} \models A$ for every $A \in \Sigma$.

One often finds the word ‘world’ (or ‘possible world’) being used for the entities in W; this use derives from our intended alethic reading of the modal language. The machinery of frames, models, and truth which we have defined is essentially an attempt to capture — by mathematical means — the view (often attributed to Leibniz) that *necessity* means *truth in all possible worlds*, and that *possibility* means *truth in some possible world*.

The truth definition stipulates that $\Diamond$ and $\Box$ check for truth not at *all* possible worlds (that is, at all elements of W) but only at R-accessible possible worlds. This may seem a weakness of the truth definition — but in fact, it is its greatest source of strength. Varying R is a mechanism which gives us a firm mathematical grip on the pre-theoretical notion of access between possible worlds. For example, by stipulating that $R = W \times W$ we can allow all worlds access to each other; this corresponds to the Leibnizian idea in its purest form. Going to the other extreme, we might stipulate that *no* world has access to any other. Between these extremes there is a wide range of options to explore. Should interworld access be reflexive? Should it be transitive? What impact do these choices have on the notions of necessity and possibility? For example, if we demand symmetry, does this justify certain principles, or rule others out?

Another philosophical issue concerns the ontological status of the states in possible worlds models. Do possible worlds exist? If they exist, are they concrete or abstract entities? Lewis [30] has been widely criticized for his concretist possible worlds realism; a well-known defender of the existence of abstract possible worlds is Plantinga [34]. Possible worlds anti-realists like Chihara [8] try to explain away metaphysical commitments of quantification over possible worlds in the metalanguage of modal logic. It seems fair to say that normally modal logicians do not feel hampered in their work by these ontological disputes.

Recall that models are composite entities consisting of a frame (our ontology) and contingent information (the valuation). We often want to ignore the effects of the valuation and get a grip on the more fundamental level of frames. The concept of *validity* lets us do this.

Definition 2 A formula A is *valid at a state* w *in a frame* $\mathcal{F}$ (notation: $\mathcal{F}, w \models A$) if A is true at w in every model $(\mathcal{F}, V)$ based on $\mathcal{F}$; A is *valid in a frame* $\mathcal{F}$ (notation: $\mathcal{F} \models A$) if it is valid at every state in $\mathcal{F}$.

For instance, $\Box(A \supset B) \supset (\Box A \supset \Box B)$ is valid on all frames. In contrast, $\Diamond\Diamond p \supset \Diamond p$ is not valid on all frames, while it is valid on all transitive frames.

What does *logical consequence* mean for modal languages? Just like we have local and global notions of truth and validity, we have two consequence relations for modal formulas. A piece of terminology: if S is a class of models, then a *model from* S is simply a model $\mathcal{M}$ in S; if S is a class of frames, then a model from S is a model based on a frame in S.

Definition 3 Let S be a class of models or a class of frames. Let Σ and A be a set of modal formula and a single formula. We say that A is a *(local) semantic consequence* of Σ over S (notation: $\Sigma \models_{\mathsf{S}} A$) if for all models $\mathcal{M}$ from S, and all states w in $\mathcal{M}$, if $\mathcal{M}, w \models \Sigma$, then $\mathcal{M}, w \models A$.

As an example, suppose that we are working with Tran, the class of frames (W, R) in which R is a transitive relation. Then $\{\Diamond\Diamond p\} \models_{\mathsf{Tran}} \Diamond p$, but $\Diamond p$ is not a local consequence of $\{\Diamond\Diamond p\}$ over the class of all frames.

Definition 4 Let A, Σ and S be as in Definition 3. Then A is a *global semantic consequence* of Σ over S (notation: $\Sigma \models^g_{\mathsf{S}} A$) if for all structures (i.e., models or frames) $\mathcal{S}$ in S, if $\mathcal{S} \models \Sigma$ then $\mathcal{S} \models A$.

The local and global notions are different, yet there is a systematic connection between them. One can show that, for Σ a set of formulas and F a class of frames, $\Sigma \models^g_{\mathsf{F}} A$ is equivalent to $\{\Box^n B \mid B \in \Sigma, n \in \omega\} \models_{\mathsf{F}} A$.

name	*formula*
D	$\Box p \supset \Diamond p$
T	$\Box p \supset p$
B	$p \supset \Box\Diamond p$
4	$\Box p \supset \Box\Box p$
5	$\Diamond A \supset \Box\Diamond A$

Table 1: Some axioms.

2.2 Completeness

During the first years after the arrival of possible worlds semantics, the topic of *axiomatic completeness* formed the bridge linking the new era with the previous syntactic era. The core notion here is that of a *normal modal logic*, which is s simply a set of formulas satisfying certain syntactic conditions. The system **K** (after Kripke) is the minimal (or 'weakest') system for reasoning about frames; stronger systems are obtained by adding extra axioms.

A *normal modal logic* Λ is a set of formulas that contains all tautologies, $\Box(p \supset q) \supset (\Box p \supset \Box q)$, and $\Diamond p \equiv \neg\Box\neg p$, and that is closed under the following three rules

- *Modus ponens*: given A and $A \supset B$, prove B.
- *Uniform substitution*: given A, prove C, where C is obtained from A by uniformly replacing proposition letters in A by arbitrary formulas.
- *Generalization*: given A, prove $\Box A$.

We write $\vdash_\Lambda A$ to denote that $A \in \Lambda$. If $\Gamma \cup \{A\}$ is a set of formulas, then A is Λ-*deducible* from Γ if either $\vdash_\Lambda A$ or there are formulas $B_1, \ldots, B_n \in \Gamma$ such that $\vdash_\Lambda (B_1 \wedge \cdots \wedge B_n) \supset A$. We call the smallest normal modal logic **K**, and a formula A is ***K**-provable* if $\vdash_{\mathbf{K}} A$. **K** is the minimal modal logic in the following sense: its axioms are all valid on all frames, and all three rules of inference preserve validity, hence all **K**-provable formulas are valid.

For many purposes **K** is too weak. For instance, if we are interested in transitive frames, we would like a proof system which reflects this. For example, we know that $\Diamond\Diamond p \supset \Diamond p$ (or equivalently $\Box p \supset \Box\Box p$) is valid on Tran, the class of all transitive frames, so we want a proof system that generates this formula. **K** does not do this, for $\Diamond\Diamond p \supset \Diamond p$ is not valid on all frames.

We can extend **K** to cope with many such semantic restrictions by adding extra axioms. Given a set of formulas Γ, we can add them as extra axioms to **K**, thus forming the axiom system $\mathbf{K\Gamma}$. Table 1 contains some familiar axioms with their traditional names.

There is a precise sense in which **K** and its extensions $\mathbf{K\Gamma}$ capture frame classes. A normal modal logic Λ is *sound* with respect to a class of frames

F if for all formulas A, $\vdash_\Lambda A$ implies $\mathcal{F} \models A$ for any $\mathcal{F} \in \mathsf{F}$. Λ is *strongly complete* with respect to F if for any set of formula $\Gamma \cup \{A\}$, if $\Gamma \models_{\mathsf{F}} A$ then $\Gamma \vdash_{\mathsf{F}} A$. Λ is *(weakly) complete* with respect to F if for any formula A, if $\mathsf{F} \models A$, then $\vdash_\Lambda A$. Table 2 lists a number of well-known modal logics together with classes of frames for which they are sound and strongly complete.

One of the most powerful methods for proving (strong) completeness results is based on *canonical models*. Given a normal logic Λ, one proves its strong completeness with respect to a class of frames F by showing that every Λ-consistent set of formulas can be satisfied in a model based on a frame in F. The canonical model method builds this model out of maximal Λ-consistent sets of formulas and uses Λ's axioms to show that the underlying frame is in F. More precisely, a set Γ is maximal Λ-consistent if it is Λ-consistent (i.e., $\Gamma \nvdash_\Lambda \bot$) and any set of formulas properly containing Γ is not Λ-consistent. By Lindenbaum's Lemma, any Λ-consistent set can be extended to a maximal consistent one. The set of maximal consistent sets forms the domain of a canonical model; the accessibility relation R in the canonical model is defined by wRv if for all formulas A, $A \in v$ implies $\Diamond A \in w$. Finally, the valuation V of the canonical model is defined by $V(p) = \{w \mid p \in w\}$.

Throughout the 1960s canonical models were the key tools used to analyze modal logics. They seem to have first been used by Makinson [31] and Cresswell [9], and in Lemmon and Scott [29] (originally written in the mid 1960s) they appear full-fledged in the form that has become standard. For a long time it was thought that every normal modal logic was complete with respect to some class of frames, and that the canonical model method could be used to prove this. The matter was resolved in 1974, when Fine [12] and Thomason [38] published examples of incomplete normal modal logics. We refer the reader to Chagrov and Zakharyaschev [7] for a modern perspective

logic	*conditions on accessibility*
K	none
KD	seriality ($\forall x \exists y\, xRy$)
KT	reflexivity
KB	symmetry
KDB	seriality, symmetry
KTB	reflexivity, symmetry
K4	transitivity
K5	Euclidicity ($\forall x \forall y \forall z\, ((xRy \wedge xRz) \supset yRz)$
KD4	seriality, transitivity
S4 (= **KT4**)	reflexivity, transitivity
S5 (= **KTB4** = **KT5**)	universal

Table 2: Some logics and their associated accessibility conditions.

and state-of-the-art account of the canonical model method.

2.3 Measuring expressive power

After the discovery of the incompleteness result, and because of an increase in interest from other disciplines to use modal logic as a *description language* for describing, e.g., process graphs or syntactic structures, attention shifted in part to *expressive power*. If we *are* using modal logic as a description language for talking about relational structures, which properties can we express? Which properties escape our description language? How can we overcome such limitations?

Before we can start answering such questions, we need to make a few things clear. First of all, recall that there are two levels at which we can use modal languages as description languages: the level of *models* and the level of *frames*, hence, the questions above can also be posed at two levels. Second, to be able to specify properties of models or frames that a modal language may or not may be able to express, we need some kind of 'background language.' For modal languages as languages for describing models we use a language of first-order logic which has unary predicate symbols P_0, P_1, P_2, ... corresponding to the proposition letters in our modal language, as well as a single binary predicate symbol R.

How are this background language and the modal language defined on page 1 related? Both can be used to talk about models of the kind used in Definition 1. For the modal language we already know this, while the only things we need to interpret the first-order language are a binary relation to interpret R (but the models of Definition 1 have that) and unary predicates to interpret P_0, P_1, P_2, ... (and, again, our models provide those, through the valuation). The modal truth definition provides the bridge between the two languages. To see this, let x be a first-order variable. The *standard translation* ST_x taking modal formulas to first-order formulas is defined as follows:

$$\begin{aligned} ST_x(p) &= Px, \\ ST_x(\neg\phi) &= \neg ST_x(\phi), \\ ST_x(\phi \vee \psi) &= ST_x(\phi) \vee ST_x(\psi), \\ ST_x(\Box\phi) &= \forall y\,(xRy \supset ST_y(\phi)), \end{aligned}$$

where y is a fresh variable (that is, a variable that has not been used so far in the translation). Note that the standard translation is nothing but a transcription of the modal truth definition in first-order logic.

As an example, $ST_x(\Diamond\Box p \supset p)$ is $\exists y\,(xRy \wedge \forall z\,(yRz \supset Pz)) \supset P(x)$.

Proposition 1 *On models, modal formulas are equivalent to their standard translations. More precisely, let A be a modal formula. Then:*

1. *For all models $\mathcal{M}$ and states w of $\mathcal{M}$: $\mathcal{M}, w \models A$ iff $\mathcal{M} \models ST_x(A)[w]$.*
2. *For all models $\mathcal{M}$, $\mathcal{M} \models A$ iff $\mathcal{M} \models \forall x\, ST_x(A)$.*

(For a first-order formula $A(x)$, the expression $\mathcal{M} \models A(x)[w]$ means that $A(x)$ is true in $\mathcal{M}$ under the assignment of w to the free variable x in $A(x)$.)

Proposition 1 may be interpreted as saying that, on models, the modal language is nothing but a fragment of the first-order language that we have specified above. But which fragment? The key notion required to answer this question is that of a *bisimulation*, introduced by van Benthem [3, 4] in the course of his work on definability and expressive power of modal logics.

Let $\mathcal{M} = (W, R, V)$ and $\mathcal{M}' = (W', R', V')$ be two models. A non-empty binary relation $Z \subseteq W \times W'$ is called a *bisimulation between* $\mathcal{M}$ and $\mathcal{M}'$ if the following conditions are satisfied:

1. If wZw' then w and w' satisfy the same proposition letters.
2. If wZw' and wRv, then there exists v' (in $\mathcal{M}'$) such that vZv' and $w'R'v'$ (the *forth condition*).
3. The converse of 2: if wZw' and $w'R'v'$, then there exists v (in $\mathcal{M}$) such that vZv' and wRv (the *back condition*).

Two states w and w' that are linked by a bisimulation are called *bisimilar*.

Proposition 2 *Modal formulas cannot distinguish between bisimilar states. That is, for all models $\mathcal{M}$ and $\mathcal{M}'$ and all states w of $\mathcal{M}$ and w' of $\mathcal{M}'$, if there is a bisimulation Z relating w to w', then $\mathcal{M}, w \models A$ iff $\mathcal{M}', w' \models A$, for all modal formulas A.*

What does Proposition 2 mean for our discussion on expressive power? By the proposition, if some property X is true of a state w and false of some w' that is bisimilar to it, then X cannot be expressed by means of a modal formula. Let us make this more concrete: consider the models $\mathcal{M}$ and $\mathcal{M}'$ shown in Figure 1. There exists a bisimulation between the models; it is given by the following relation Z: $Z = \{(1, a), (2, b), (2, c), (3, d), (4, e), (5, e)\}$. Condition 1 of the definition of a bisimulation is obviously satisfied: Z-related states make the same propositional letters true. Moreover, the back and forth conditions are satisfied too: any move in $\mathcal{M}$ can be matched by a similar move in $\mathcal{M}'$, and conversely.
There are some obvious differences between, for instance, the state 3 in $\mathcal{M}$ and the state d in $\mathcal{M}'$, despite the fact that they are bisimilar. For instance, the property $\exists y \exists z\, (xRy \wedge xRz \wedge y \neq z \wedge P(y) \wedge P(z))$ is true of 3 in $\mathcal{M}$ but not of d in $\mathcal{M}'$. Hence, by Proposition 2, this property is not expressible by a modal formula.

But we can get more out of bisimulations. By a famous result due to van Benthem, the inability to distinguish between bisimilar states is characteristic of the modal fragment:

Theorem 1 (van Benthem Characterization Theorem) *Let $A(x)$ be a first-order formula (over a vocabulary consisting of R, P_0, P_1, P_2, ...). Then $A(x)$ is equivalent to the standard translation of a modal formula iff it cannot distinguish between bisimilar states.*

The above result was first proved by van Benthem in his PhD thesis [3]; see also [4]. Analogous bisimulation-based characterizations have since been given for a wide variety of modal and modal-like languages; consult [5] for an overview.

We now turn to a brief discussion of the expressive power of the modal language as a language for talking about frames. We start by explaining why frame definability is intrinsically second-order, and give examples of frame classes that are modally definable but not first-order definable. Recall that validity is defined as quantifying over all states of the domain of a frame and over all possible valuations. But a valuation assigns a *subset* of a frame to each proposition letter, and this means that when we quantify across all valuations, we are implicitly quantifying across all subsets of the frame. We can make this more precise in the following manner: we saw that at the level of models, the modal language can be translated in a truth-preserving way into a first-order language — but we can view the predicate symbols P_0, P_1, P_2, ... that correspond to the proposition letters p_0, p_1, p_2, ... as monadic second-order variables that we can quantify over. If we do this, we are in effect viewing the standard translation as a way of translating into a second-order language with a binary relation symbol, and monadic predicate variables P_0, P_1, P_2, ... This leads to the following result:

Proposition 3 *Let A be a modal formula. Then the following holds for any frame $\mathcal{F}$ and any state w of $\mathcal{F}$:*

1. $\mathcal{F}, w \models A$ *iff* $\mathcal{F} \models \forall P_1 \ldots \forall P_n \, ST_x(A)[w]$
2. $\mathcal{F} \models A$ *iff* $\mathcal{F} \models \forall P_1 \ldots \forall P_n \, ST_x(A)$

As a concrete example, it can be shown that a formula as simple as the McKinsey formula $\Diamond\Box p \supset \Box\Diamond p$ is essentially a second-order formula when

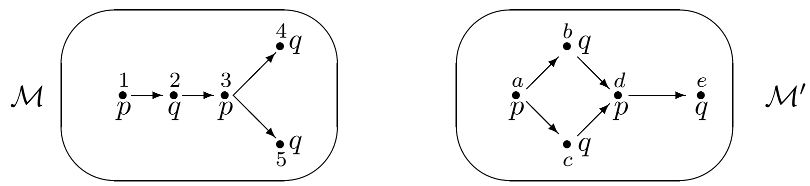

Figure 1: Bisimilar models.

interpreted on frames: there is an uncountable frame $\mathcal{F}$ on which the McKinsey formula is valid, while it is *in*valid on each of $\mathcal{F}$s countable elementary subframes, thus showing that the McKinsey formula violates the downward Löwenheim-Skolem Theorem, one of the essential model-theoretic properties of first-order logic.

There are many modal formulas that define first-order conditions on frames. Tables 1 and 2 provide examples. Given that we have just seen that frame definability is a second-order notion, this is a surprising result. It turns out that in many cases the (often difficult to decipher) second-order condition produced by second-order translation is equivalent to a much simpler first-order condition. There exists an algorithm, called the *Sahlqvist-van Benthem* algorithm, that computes a corresponding first-order condition for a large class of modal formulas; this is the celebrated *Sahlqvist Correspondence Theorem.*

To be able to define the class of formulas for which the Sahlqvist-van Benthem algorithm works, we need the following shorthand: a *boxed atom* is a formula of the form $\Box \cdots \Box p$; in the case where the number of boxes preceding p is 0, the boxed atom $\Box \cdots \Box p$ is just the proposition letter p. Next, a *negative* formula is one in which all occurrences of proposition letters are in the scope of an odd number of negation signs. Furthermore, a *Sahlqvist antecedent* is a formula built up from $\top$, $\bot$, boxed atoms, and negative formulas, using $\wedge$, $\vee$ and $\Diamond$. A *Sahlqvist implication* is an implication $A \supset B$ in which B is positive and A is a Sahlqvist antecedent. Finally, then, a *Sahlqvist formula* is a formula that is built up from Sahlqvist implications by freely applying boxes and conjunctions, and by applying disjunctions only between formulas that do not share any proposition letters.

Examples of Sahlqvist formulas include $\Box(p \supset \Diamond p)$, and the axioms D, T, B, 4, and 5 from Table 1. Typically forbidden combinations in Sahlqvist antecedents are 'boxes over disjunctions,' and 'boxes over diamonds,' as illustrated by the McKinsey formula.

Theorem 2 (Sahqlvist Correspondence Theorem) *Let A be a Sahlqvist formula. Then, on frames, A is equivalent to a first-order condition $C_A(x)$ that is effectively computable from A.*

The key idea underlying the proof of the above result is the following: strip off the initial block of monadic second-order universal quantifiers in $\forall P_1 \ldots \forall P_n \, ST_x(A)$, thus reducing it to a first-order formula. The obvious way of getting rid of universal quantifiers is to perform universal instantiation, but the key point underlying the proof of the Sahqlvist Correspondence Theorem is that, in the case of Sahlqvist formulas, instantiations can be chosen in such a way that the resulting first-order formula is equivalent to (and not just implied by) the original second-order formula.

To illustrate the point, consider the Sahlqvist formula $(p \wedge \Diamond\neg p) \supset \Diamond p$. Its second-order translation is

$$\forall P\,(Px \wedge \exists y\,(Rxy \wedge \neg Py) \supset \exists z\,(Rxz \wedge Pz)).$$

Pulling out the existential quantifier produces

$$\forall P \forall y\,(Px \wedge Rxy \wedge \neg Py \supset \exists z\,(Rxz \wedge Pz)),$$

and moving the negative part $\neg Py$ to the consequent we get

$$\forall P \forall y\,(Px \wedge Rxy \supset Py \vee \exists z\,(Rxz \wedge Pz)). \qquad (1)$$

The minimal instantiation to make Px true is one that assigns P to an object u iff $u = x$. After instantiation we obtain

$$\forall y\,(Rxy \supset y = x \vee \exists z\,(Rxz \wedge z = x)),$$

and it can be shown that this is actually equivalent to (1). The latter can of course be simplified to $\forall y\,(Rxy \wedge x \neq y \supset Rxx)$.

The Sahlqvist Correspondence Theorem comes together with a Sahlqvist Completeness Theorem: not only does every Sahlqvist formula correspond to a first-order property of frames, but when we use one as an axiom in a normal modal logic, that logic is guaranteed to be *complete* with respect to the class of frames defined by the first-order property! Moreover, the completeness result can be proved using the canonical model method; see [5] for details.

To conclude our discussion of Sahlqvist formulas, we want to mention a result due to Kracht, who has isolated the first-order formulas that are the correspondents of Sahlqvist formulas in [24, 26], as an application of his so-called calculus of internal describability. Unfortunately, the details are too technical to be included here; see also [5].

While Kracht's result gives us insight into the first-order frame properties definable by means of Sahlqvist formulas, it does not provide us with a complete description of the modally definable properties. For this, we have to turn to the Goldblatt-Thomason Theorem. The result characterizes the expressive power of modal languages on frames in terms of four fundamental frame constructions: *disjoint unions*, *generated subframes*, *bounded morphic images*, and *ultrafilter extensions*. Here, the disjoint union $\mathcal{F}$ of two frames $\mathcal{F}_1$ and $\mathcal{F}_2$ simply has the disjoint union of the domains of $\mathcal{F}_1$ and $\mathcal{F}_2$ as its domain, while its relation is the disjoint union of the relations for $\mathcal{F}_1$ and $\mathcal{F}_2$. Moreover, $\mathcal{F}_1 = (W_1, R_1)$ is a generated subframe of $\mathcal{F}_2 = (W_2, R_2)$ if W_1 is a subset of W_2 that is closed under the addition of R_2-related states, while R_1 is simply the restriction of R_2 to W_1. A bounded morphism is nothing but a functional version of our earlier notion of bisimulation, adapted to the case of frames. And, finally, the ultrafilter extension of a frame is a kind of completion of the original frame; they are built by using the ultrafilters over a given frame as the states of a new frame, and defining an appropriate relation between them; see [5] for formal and informal explanations.

Theorem 3 (Goldblatt-Thomason Theorem) *Let K be a class of frames that is defined by a first-order sentence. That is, let K be such that for some first-order sentence A, we have that, for all frames $\mathcal{F}$, $\mathcal{F} \in \mathsf{K}$ iff $\mathcal{F}$ satisfies A. Then K is definable by means a modal formula iff it is closed under bounded morphic images, generated subframes, disjoint unions while it reflects ultrafilter extensions in the sense $\mathcal{F} \in \mathsf{K}$ whenever the ultrafilter extension $\mathcal{F}$ is in K.*

The Goldblatt-Thomason Theorem was actually proved by Goldblatt. His original result was stronger than the one we have given, applying to any frame class that is closed under elementary equivalence; this result was published in a joint paper [19] with S.K. Thomason.

3 Proof Theory

Although modern alethic modal logic started as a syntactic enterprise, its proof theory was somewhat neglected after the advent of possible worlds semantics. An exception is the development of semantic tableau calculi for modal logic. Tableau proof systems amount to rules for the construction of countermodels and take into account the relational patterns of possible worlds models. We will first consider semantic tableaux and then 'display logic', a generalization of Gentzen's sequent calculus based on the idea of residuation and Galois connection.

3.1 Tableau calculi

Tableau calculi incorporating the accessibility relation of possible worlds models were first introduced by Kripke [27] and were later 'linearized' by various authors, notably Fitting [13, 14] and Mints [33]. The basic declarative unit of these calculi is not just a formula A, but rather a *formula plus label* (σ, A). In general, the label σ is a non-empty finite sequence of positive integers. A simplification is possible for **S5**. Since **S5** is characterized by the class of all frames with a universal accessibility relation R, R can be neglected, and the label σ may just be a single positive integer. A comprehensive survey on tableau methods for modal and tense logics is [21]. The use of labels allows one to formulate tableau calculi for certain extensions of the minimal normal modal logic **K** by imposing constraints on accessibility and on occurrences and the shape of labels on tableau branches.

A Gentzen sequent is an expression $\Delta \rightarrow \Gamma$, where Δ and Γ are finite sets of formulas, and $\Delta \rightarrow \Gamma$ is to be understood as the claim that $\bigwedge \Delta \supset \bigvee \Gamma$ is provable. In (extensions of) classical logic, the latter formula is valid iff the set $\{A \mid A \in \Delta\} \cup \{\neg B \mid B \in \Gamma\}$ fails to be satisfiable. Rules for manipulating the sequent $\Delta \rightarrow \Gamma$ can therefore also be stated as rules for manipulating the finite set $\{A \mid A \in \Delta\} \cup \{\neg B \mid B \in \Gamma\}$. Although tableau calculi are often presented using the set notation, we here prefer a sequent notation. We will use bold letters $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$ (possibly with primes or subscripts) to denote arbitrary finite sets of labelled formulas. A *sequent* is an expression of the form $\mathbf{X} \rightarrow \mathbf{Y}$, where $\mathbf{X}$ is called the *antecedent* and $\mathbf{Y}$ is called the *succedent* of this sequent. We use s, s_1, s_2, ... to denote sequents and the 'turnstile' $\vdash$ to denote derivability between single sequents and finite sets of sequents.

Tableau calculi are given by (finite) sets of derivation rules of the form $\mathsf{s} \vdash \mathsf{s}_1, \ldots, \mathsf{s}_n$. A *tableau* for a given sequent s is a tree of sequents rooted in s, such that every node of the tree is an instantiation of one of the derivation rules of the tableau calculus under consideration. A tableau for s is *closed* if every leaf of any branch of the tableau has the form $(\sigma, B) \rightarrow (\sigma, B)$. We assume a binary relation of 'accessibility' between labels. This relation may satisfy certain conditions, and a number of such conditions is defined in Table 3.

The logic $\mathbf{K}$ and various extensions of it that can be dealt with by means of labelled tableaux, require certain properties of accessibility between labels. Table 4 lists such systems together with the required properties of accessibility between labels. A label τ occurring on a tableau branch is said to be a *simple, unrestricted extension* of a label σ iff (i) τ is the result of extending σ on the right with a single positive integer and (ii) τ is not an initial segment of any label occurring on the branch. The label τ is *available* on a branch if it occurs on that branch. Right and left introduction rules for $\Box$ can now be stated in such a way that variations among the systems listed in Table 4 can be accounted for by side conditions on the left rule (cf. [14, p. 402]). These rules are stated in Table 5 together with the tableau rules for disjunction and negation.

name of condition	*definition*
general	for every n, σn is accessible from σ
symmetry	for every n, σ is accessible from σn
reflexivity	σ is accessible from σ
transitivity	if σ is a proper initial segment of τ, then τ is accessible from σ
universal	any label is accessible from any label

Table 3: Conditions on accessibility.

logic	*conditions on accessibility*
K, **KD**	general
KT	general, reflexivity
KB, **KDB**	general, symmetry
KTB	general, reflexivity, symmetry
K4, **KD4**	general, transitivity
S4 (= **KT4**)	general, reflexivity, transitivity
S5 (= **KTB4**)	universal

Table 4: Some logics and their associated label accessibility conditions.

name	*rule*
L$\vee$	$\mathbf{X}, (\sigma, A \vee B) \to \mathbf{Y} \vdash \mathbf{X}, (\sigma, A) \to \mathbf{Y} \quad \mathbf{X}, (\sigma, B) \to \mathbf{Y}$
R$\vee$	$\mathbf{X} \to (\sigma, A \vee B), \mathbf{Y} \vdash \mathbf{X} \to (\sigma, A), (\sigma, B), \mathbf{Y}$
L$\neg$	$\mathbf{X}, (\sigma, \neg A) \to \mathbf{Y} \vdash \mathbf{X} \to (\sigma, A), \mathbf{Y}$
R$\neg$	$\mathbf{X} \to (\sigma, \neg A), \mathbf{Y} \vdash \mathbf{X}, (\sigma, A) \to \mathbf{Y}$
L$\Box$	$\mathbf{X}, (\sigma, \Box A) \to \mathbf{Y} \vdash \mathbf{X}, (\tau, A) \to \mathbf{Y}$ for any τ accessible from σ provided (i) for **K**, **KB**, and **K4**, τ must be available on the branch; (ii) for **KD**, **KT**, **KDB**, **KTB**, **KD4**, **S4**, and **S5**, τ must either be available on the branch or be a simple, unrestricted extension of σ
R$\Box$	$\mathbf{X} \to (\sigma, \Box A), \mathbf{Y} \vdash \mathbf{X} \to (\tau, A), \mathbf{Y}$ provided τ is a simple, unrestricted extension of σ

Table 5: Tableau rules.

For every logic Λ from Table 4, let $T\Lambda$ be its tableau calculus. Here is an example (1) of a closed tableau for $\emptyset \to (1, \neg\Box A \vee \Box\Box A)$ in T**K4** and an example (2) of a closed tableau for $\emptyset \to (1, \neg A \vee \Box\neg\Box\neg A)$ in T**KB**.

(1)

$$\emptyset \to (1, \neg\Box A \vee \Box\Box A)$$
$$\emptyset \to (1, \neg\Box A),\ (1, \Box\Box A)$$
$$(1, \Box A) \to (1, \Box\Box A)$$
$$(1, \Box A) \to (\langle 1, 2\rangle, \Box A)$$
$$(1, \Box A) \to (\langle 1, 2, 3\rangle, A)$$
$$(\langle 1, 2, 3, \rangle, A) \to (\langle 1, 2, 3\rangle, A)$$

(2)

$$\emptyset \to (1, \neg A \vee \Box\neg\Box\neg A)$$
$$\emptyset \to (1, \neg A),\ (1, \Box\neg\Box\neg A)$$
$$(1, A) \to (1, \Box\neg\Box\neg A)$$
$$(1, A) \to (\langle 1, 2\rangle, \neg\Box\neg A)$$
$$(1, A),\ (\langle 1, 2\rangle, \Box\neg A) \to \emptyset$$
$$(1, A),\ (1, \neg A) \to \emptyset$$
$$(1, A) \to (1, A)$$

Theorem 4 *A modal formula A is a theorem of a logic Λ from Table 4 iff there is a closed tableau for $\emptyset \to (1, A)$ in $T\Lambda$.*

3.2 Display calculi

The display calculus [1] is a generalization of Gentzen's sequent calculus. We will present display logic only to the extent needed to treat normal modal logics. A more comprehensive presentation of display logic and its application to modal and non-classical logics can be found in [1, 2, 20, 25, 36, 39]. The modal display calculus is based on the observation that the operators $\blacklozenge$ ("sometimes in the past", i.e., the possibility operator with respect to the inverse $R^{\smile}$ of the accessibility relation R) and $\Box$ form a residuated pair. The following definition is taken from Dunn [11, p. 32]:

Definition 5 Let $\mathcal{A} = (\mathsf{A}, \leq)$ and $\mathcal{B} = (\mathsf{B}, \leq')$ be partially ordered sets with functions f: $\mathsf{A} \longrightarrow \mathsf{B}$ and g: $\mathsf{B} \longrightarrow \mathsf{A}$. The pair (f, g) is called

- *residuated* iff ($fa \leq' b$ iff $a \leq gb$);
- a *Galois connection* iff ($b \leq' fa$ iff $a \leq gb$);
- a *dual Galois connection* iff ($fa \leq' b$ iff $gb \leq a$);
- a *dual residuated pair* iff ($b \leq' fa$ iff $gb \leq a$).

Obviously, $(\blacklozenge, \Box)$ forms a residuated pair with respect to the (local) semantic consequence relation $\models$ with respect to classes of Kripke frames. These ideas of residuation and Galois connection can be generalized, but for our purposes we have all we need to formulate introduction sequent rules for the modal operators. The polyvalent comma as a structure connective in Gentzen's sequent calculus is replaced by a number of structure connectives: $\boldsymbol{I}$ (nullary), $*$ (unary), $\bullet$ (unary), $\circ$ (binary). Every formula A is a structure, and we will use X, Y, and Z as variables for structures. The structures are defined by:

$$X ::= A \mid \boldsymbol{I} \mid *X \mid \bullet X \mid X \circ Y.$$

A *display sequent* is an expression $X \rightarrow Y$; X is called the *antecedent* and Y the *succedent* of $X \rightarrow Y$. The intended meaning of the structure connectives can be made explicit by a translation $\mathtt{t}(X \rightarrow Y) := \mathtt{t}_1(X) \supset \mathtt{t}_2(Y)$ of sequents into formulas, where $\mathtt{t}_i(A) = A$ $(i = 1, 2)$, and:

$$\begin{array}{rclrcl}
\mathtt{t}_1(\boldsymbol{I}) & = & \top & \mathtt{t}_2(\boldsymbol{I}) & = & \bot \\
\mathtt{t}_1(*X) & = & \neg\mathtt{t}_2(X) & \mathtt{t}_2(*X) & = & \neg\mathtt{t}_1(X) \\
\mathtt{t}_1(\bullet X) & = & \blacklozenge\mathtt{t}_1(X) & \mathtt{t}_2(\bullet X) & = & \Box\mathtt{t}_2(X) \\
\mathtt{t}_1(X \circ Y) & = & \mathtt{t}_1(X) \wedge \mathtt{t}_1(Y) & \mathtt{t}_2(X \circ Y) & = & \mathtt{t}_2(X) \vee \mathtt{t}_2(Y).
\end{array}$$

Under the $\mathtt{t}$-translation, the following basic structural rules are valid in every normal modal logic:

(1) $X \circ Y \rightarrow Z \dashv\vdash X \rightarrow Z \circ *Y \dashv\vdash Y \rightarrow *X \circ Z$

(2) $X \rightarrow Y \circ Z \dashv\vdash X \circ *Z \rightarrow Y \dashv\vdash *Y \circ X \rightarrow Z$

(3) $X \rightarrow Y \dashv\vdash *Y \rightarrow *X \dashv\vdash X \rightarrow **Y$

name	*rule*
$(\rightarrow \neg)$	$X \rightarrow *A \vdash X \rightarrow \neg A$
$(\neg \rightarrow)$	$*A \rightarrow X \vdash \neg A \rightarrow X$
$(\rightarrow \vee)$	$X \rightarrow A \circ B \vdash X \rightarrow A \vee B$
$(\vee \rightarrow)$	$A \rightarrow X \quad B \rightarrow Y \vdash A \vee B \rightarrow X \circ Y$
$(\rightarrow \Box)$	$\bullet X \rightarrow A \vdash X \rightarrow \Box A$
$(\Box \rightarrow)$	$A \rightarrow X \vdash \Box A \rightarrow \bullet X$

Table 6: Introduction rules for the logical operations.

(4) $X \rightarrow \bullet Y \dashv\vdash \bullet X \rightarrow Y$.

Here, $X_1 \rightarrow Y_1 \dashv\vdash X_2 \rightarrow Y_2$ is an abbreviation of $X_1 \rightarrow Y_1 \vdash X_2 \rightarrow Y_2$ and $X_2 \rightarrow Y_2 \vdash X_1 \rightarrow Y_1$. If two sequents are interderivable by means of (1)–(4), they are said to be *structurally* or *display equivalent*. The name 'display logic' is due to the fact that any substructure of a given display sequent s may be displayed as the entire antecedent or succedent of a structurally equivalent sequent s'. In order to state this fact precisely, we define the notion of antecedent and succedent part of a sequent. An occurrence of a substructure in a given structure is called positive (negative), if it is in the scope of an even (odd) number of $*$'s. An antecedent (succedent) part of a sequent $X \rightarrow Y$ is a positive occurrence of a substructure of X or a negative occurrence of a substructure of Y (a negative occurrence of a substructure of X or a positive occurrence of a substructure of Y).

Theorem 5 (Display Theorem, Belnap [1]) *For every display sequent* s *and every antecedent (succedent) part* X *of* s *there exists a display sequent* s' *structurally equivalent to* s *such that* X *is the entire antecedent (succedent) of* s'.

The structure connectives $*$, $\boldsymbol{I}$ and $\circ$ give rise to introduction rules for the Boolean connectives, and $\bullet$ permits formulating introduction rules for $\Box$. These introduction rules are presented in Table 6. Table 7 collects

name	*rule*
(I)	$X \rightarrow Z \dashv\vdash \boldsymbol{I} \circ X \rightarrow Z, \qquad X \rightarrow Z \dashv\vdash X \rightarrow \boldsymbol{I} \circ Z$
(A)	$X_1 \circ (X_2 \circ X_3) \rightarrow Z \dashv\vdash (X_1 \circ X_2) \circ X_3 \rightarrow Z$
(P)	$X_1 \circ X_2 \rightarrow Z \vdash X_2 \circ X_1 \rightarrow Z$
(C)	$X \circ X \rightarrow Z \vdash X \rightarrow Z$
(M)	$X \rightarrow Z \vdash X \circ Y \rightarrow Z$
(MN)	$\boldsymbol{I} \rightarrow X \vdash \bullet \boldsymbol{I} \rightarrow X$

Table 7: Additional structural rules.

$$A \rightarrow A$$
$$\Box A \rightarrow \bullet A$$
$$\Box(\neg A \vee B) \circ \Box A \rightarrow \bullet A \quad (bs)$$
$$* A \rightarrow * \bullet (\Box(\neg A \vee B) \circ \Box A)$$
$$\neg A \rightarrow * \bullet (\Box(\neg A \vee B) \circ \Box A) \qquad B \rightarrow B$$
$$\neg A \vee B \rightarrow * \bullet (\Box(\neg A \vee B) \circ \Box A) \circ B$$
$$\Box(\neg A \vee B) \rightarrow \bullet(* \bullet (\Box(\neg A \vee B) \circ \Box A)) \circ B$$
$$\Box(\neg A \vee B) \circ \Box A \rightarrow \bullet(* \bullet (\Box(\neg A \vee B) \circ \Box A)) \circ B \quad (bs)$$
$$\bullet(\Box(\neg A \vee B) \circ \Box A) \circ \bullet(\Box(\neg A \vee B) \circ \Box A) \rightarrow B$$
$$\bullet(\Box(\neg A \vee B) \circ \Box A) \rightarrow B$$
$$\Box(\neg A \vee B) \circ \Box A \rightarrow \Box B \quad (bs)$$
$$* \Box B \circ \Box(\neg A \vee B) \rightarrow * \Box A$$
$$* \Box B \circ \Box(\neg A \vee B) \rightarrow \neg \Box A \quad (bs)$$
$$\Box(\neg A \vee B) \rightarrow \neg \Box A \vee \Box B$$

Figure 2: A derivation in **DK**.

further structural rules that together with the introduction rules ensure the classical and normal modal behavior of the logical operations. A richer inventory of structural rules (and another choice of structure connectives) is called for in display calculi for substructural logics, see [20]. In addition to structural rules and introduction rules, every display calculus contains two distinguished *logical* (structural) rules, namely identity for atoms and cut:

(identity) $\vdash p \rightarrow p$
(cut) $\quad X \rightarrow A \quad A \rightarrow Y \vdash X \rightarrow Y$

It can be shown by induction on formulas A that $\vdash A \rightarrow A$. The display calculus **DK** consists of (id), (cut), the basic and additional structural rules and the introduction rules for $\neg$, $\vee$, and $\Box$. As an example, Figure 2 depicts a cut-free derivation of $\Box(\neg A \vee B) \rightarrow \neg\Box A \vee \Box B$, where (bs) indicates the repeated application of some basic structural rules.
Using induction on the complexity of X, it can be shown that in every extension of **DK** by structural rules, $\vdash X \rightarrow \mathtt{t}_1(X)$ and $\vdash \mathtt{t}_2(X) \rightarrow X$. This observation is used in the proof of the characterization theorem.

Theorem 6 *In* ***DK****,* $\vdash X \rightarrow Y$ *iff* $\mathtt{t}_1(X) \supset \mathtt{t}_2(Y)$ *is provable in* ***K****.*

A display sequent system is said to be a *proper* display calculus, if it satisfies certain conditions $C1$–$C8$ first stated by Belnap [1]. A logic is said to be *properly displayable*, if it can be presented as a proper display calculus. Every proper display calculus enjoys cut-elimination [1] and even strong cut-elimination [39]. In this case, strong cut-elimination means that there is a set of reduction steps for turning a given sequent proof into a cut-free

schema	*structural rule*
D	$* \bullet * \boldsymbol{I} \rightarrow Y \vdash \boldsymbol{I} \rightarrow Y$
T	$X \rightarrow \bullet Y \vdash X \rightarrow Y$
4	$X \rightarrow \bullet Y \vdash X \rightarrow \bullet \bullet Y$
B	$* \bullet *(X \circ * \bullet * Y) \rightarrow Z \vdash Y \circ * \bullet * X \rightarrow Z$
5	$* \bullet * X \rightarrow Y \vdash \bullet * \bullet * X \rightarrow Y$

Table 8: Axiom schemata and corresponding structural rules.

proof of the same sequent such that — *modulo certain mild restrictions* — every sufficiently long sequence of applications of these reduction steps to a proof Π will return a cut-free proof Π' of the same sequent. The class of all properly displayable extensions of the smallest normal temporal logic has been characterized by Kracht [25].

Here we will just consider display calculi for extensions of **K** by the familiar and important axiom schemata D, T, 4, B and 5 that correspond to the seriality, reflexivity, transitivity, symmetry, and Euclidicity, respectively, of the accessibility relation R. It turns out that these axiom schemata can be captured by the purely structural rules stated in Table 8. Let $\theta \subseteq \{D, T, 4, B, 5\}$, $\overline{\theta} = \{\alpha' \mid \alpha \in \theta\}$. Let $\mathbf{K}\theta$ be the result of adding the axiom schemata from θ to **K**, and let $\mathbf{DK}\theta'$ be the result of adding the structural rules from θ' to **DK**.

Theorem 7 *In* $\boldsymbol{DK}\theta$, $\vdash X \rightarrow Y$ *iff* $\mathtt{t}_1(X) \supset \mathtt{t}_2(Y)$ *is provable in* $\boldsymbol{K}\theta'$.

4 Modal Predicate Logic

While propositional modal logic has become a highly developed discipline with a broad spectrum of choices as regards expressive power and reasoning methods, in some cases the added modeling power of modal *predicate* logic is called for. Below we briefly discuss some of the philosophical and mathematical issues involved with this choice.

In modal predicate logic there are various junctions where metaphysics, philosophy of language and formal logic meet. Let $\mathcal{F} = (W, R)$ be a frame. If to every state $s \in W$ a domain $D = d(w)$ is associated, there are at least the following, well-known options:

1. $(\forall s \in W)$, $d(s) \neq \emptyset$; (varying domains);
2. $(\forall s, t \in W)$, $d(s) \neq \emptyset$ and if sRt, then $d(s) \subseteq d(t)$ (increasing domains);
3. $(\forall s, t \in W)$, $d(s) \neq \emptyset$ and $d(s) = d(t)$ (constant domains).

Is every individual present in every state? What are the effects a state transition can have on a domain? It seems natural to assume that if sRt, individuals not already present in s may appear in t or individuals present in s

may disappear in t. With a fixed set of individual constants, the assumptions of varying and increasing domains permit non-designating ground (that is, variable-free) terms. In addition to the semantical problem of interpreting non-designating ground terms and formulas containing such terms, the metaphysical question arises, whether an individual may or not possess properties in a state where the individual does not exist. The assumption of constant domains corresponds to the validity of the Barcan formula $\forall x \Box A \supset \Box \forall x A$ and the assumption of increasing domains corresponds to the validity of the converse Barcan formula $\Box \forall x A \supset \forall x \Box A$. We refer to the recent [15] for an overview of discussions of these and related matters.

There is a whole web of mathematical questions related the Barcan formula and its variations. As to proof-theoretical aspects, the standard ordinary sequent calculus for **K** uses Gentzen sequents $\Delta \rightarrow \Gamma$ and comprises just one introduction rule for $\Box$, namely

$$\Delta \rightarrow A \vdash \{\Box B \mid B \in \Delta\} \rightarrow \Box A$$

If this calculus is enlarged by the familiar introduction rules for the universal quantifier, the Barcan formula and its converse are derivable. This fact supports the idea that modal logic requires a generalized notion of sequent.

It has often been observed that $\Box$ is a universal quantifier over possible worlds in the metalanguage of modal logic. In display logic, a universal quantifier prefix $\forall x$ can be treated like the necessity operator, by associating with $\forall x$ a structure operation $\bullet_x$ and a binary relation R_x such that in succedent position $\bullet_x A$ is interpreted as $\forall x A$ and in antecedent position as $\exists x\breve{}$, the 'possibility' operator with respect to the converse relation $R_x\breve{}$ of R_x. The Barcan formula and its converse then correspond to additional structural rules, for details see [39]:

$$X \rightarrow \bullet_x \bullet Y \vdash X \rightarrow \bullet \bullet_x Y; \qquad X \rightarrow \bullet \bullet_x Y \vdash X \rightarrow \bullet_x \bullet Y.$$

Tableaux calculi for modal predicate logics with and without the Barcan formula can be found in [32].

Just like the identity of individuals gives rise to many philosophical questions in modal predicate logic, it also gives rise to many deep mathematical questions. As a result, various alternative semantic frameworks have been developed for modal predicate logic during the 1990s, including the Kripke bundles of Shehtman and Skvortsov [37] and the category-theoretic semantics proposed by Ghilardi [16]

The notion of (axiomatic) completeness is another source of interesting mathematical questions in modal predicate logic. It turns out that the minimal predicate logical extension of many well-behaved and complete propositional modal logics need not be complete. The main (negative) result in this area is that among the extensions of **S4**, propositional modal logics **L** whose minimal predicate logical extension is complete must have either $\mathbf{L} \supseteq \mathbf{S5}$ or

L $\subseteq$ **S4.3**. This excludes completeness results for predicate logical extensions for logics such as **S4.1** and **S4.3Grz**. Positive completeness results are known only for some boundary cases: the predicate logical extensions of **S4**, **S4.2**, **S4.3**, and **S5** and its extensions; see [10] for a recent overview.

Still further mathematical questions come up in the search for algorithmically well-behaved fragments of modal predicate logics; very powerful results were recently obtained by Hodkinson, Wolter, and Zakharyaschev [22].

5 Further Readings

We conclude this chapter with some pointers to the literature on modal logic. First, details on the history of modern modal logic are available, for instance, in [5, 6, 18, 40]. Second, there are several survey papers in recent and not so recent handbooks that can serve as valuable starting points for further studies; these include [6]. Third, there is a broad range of modern textbooks on modal logic, ranging from philosophically oriented [17, 23] to more mathematically inclined [5, 35]. Finally, the Advances in Modal Logic initiative, with its accompanying workshops, volumes, and web site at `www.aiml.net` is a rich source of information.

References

[1] N.D. Belnap. Display logic. *Journal of Philosophical Logic*, 11:375–417, 1982. Reprinted as §62 of A.R. Anderson, N.D. Belnap, and J.M. Dunn, *Entailment: the logic of relevance and necessity.* Vol. 2, Princeton University Press, 1992.

[2] N.D. Belnap. Linear logic displayed. *Notre Dame Journal of Formal Logic*, 31:14–25, 1990.

[3] J. van Benthem. *Modal Correspondence Theory.* PhD thesis, Mathematisch Instituut & Instituut voor Grondslagenonderzoek, University of Amsterdam, 1976.

[4] J. van Benthem. *Modal Logic and Classical Logic.* Bibliopolis, 1983.

[5] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic.* Cambridge University Press, 2001. See also `http://www.mlbook.org`.

[6] R. Bull and K. Segerberg. Basic modal logic. In D.M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, Vol. II*, pages 1–88. Reidel, 1984.

[7] A. Chagrov and M. Zakharyaschev. *Modal Logic*, volume 35 of *Oxford Logic Guides.* Oxford University Press, Oxford, 1997.

[8] C. Chihara. *The Worlds of Possibility.* Oxford University Press, 1998.

[9] M.J. Cresswell. A Henkin completeness theorem for T. *Notre Dame Journal of Formal Logic*, 8:186–90, 1967.

[10] M.J. Cresswell. How to complete some modal predicate logics. In *Advances in Modal Logic. Volume 2*, 2001.

[11] J. M. Dunn. Gaggle theory: An abstraction of galois connections and residuation with applications to negation and various logical operations. In J. van Eijck, editor, *Proceedings JELIA 1990*, pages 31–51. Springer, 1990.

[12] K. Fine. An incomplete logic containing $S4$. *Theoria*, 40:23–29, 1974.

[13] M. Fitting. *Proof Methods for Modal and Intuitionistic Logics*. Reidel, 1983.

[14] M. Fitting. Basic modal logic. In D.M. Gabbay et al., editor, *Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 1*, pages 365–448. Oxford University Press, 1993.

[15] M. Fitting and R.L. Mendelsohn. *First-order Modal Logic*. Kluwer Academic Publishers, 1998.

[16] S. Ghilardi. Incompleteness results in Kripke semantics. *Journal of Symbolic Logic*, 56:517–538, 1991.

[17] R. Girle. *Modal Logics and Philosophy*. Acumen, 2000.

[18] R. Goldblatt. Mathematical modal logic: a view of its evolution, 2000. To appear. Draft available at `http://www.vuw.ac.nz/~rob`.

[19] R.I. Goldblatt and S.K. Thomason. Axiomatic classes in propositional modal logic. In J. Crossley, editor, *Algebra and Logic*, pages 163–173. Springer, 1974.

[20] R. Goré. Substructural logics on display. *Logic Journal of the IGPL*, 6:451–504, 1998.

[21] R. Goré. Tableau methods for modal and temporal logics. In M. D'Agostino, D.M. Gabbay, R. Hähnle, and J. Posegga, editors, *Handbook of Tableau Methods*. Kluwer Academic Publishers, 1999.

[22] I.M. Hodkinson, F. Wolter, and M. Zakharyaschev. Decidable fragments of first-order temporal logics. *Anals of Pure and Applied Logic*, 106:85–134, 2000.

[23] G. Hughes and M.J. Cresswell. *A New Introduction to Modal Logic*. Routledge, 1996.

[24] M. Kracht. How completeness and correspondence theory got married. In M. de Rijke, editor, *Diamonds and Defaults*, pages 175–214, 1993.

[25] M. Kracht. Power and weakness of the modal display calculus. In H. Wansing, editor, *Proof Theory of Modal Logic*, pages 93–121. Kluwer Academic Publishers, 1996.

[26] M. Kracht. *Tools and Techniques in Modal Logic*. Number 142 in Studies in Logic. Elsevier, Amsterdam, 1999.

[27] S. Kripke. Semantical analysis of modal logic I: Normal modal propositional calculi. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 9:67–96, 1963.

[28] A.R. Lacey. *A Dictionary of Philosophy*. Routledge & Kegan Paul, 1976.

[29] E.J. Lemmon and D.S. Scott. *The "Lemmon Notes": An Introduction to Modal Logic*. Blackwell, Oxford, 1977.

[30] D. Lewis. *On the Plurality of Worlds*. Blackwell, 1986.

[31] D.C. Makinson. On some completeness theorems in modal logic. *Zeitschrift für mathemathische Logik und Grundlagen der Mathematik*, 12:379–84, 1966.

[32] M. Cialdea Mayer and S. Cerrito. Variants of first-order modal logic. In R. Dyckhoff, editor, *Proceedings Tableaux 2000*. Springer, 2000.

[33] G. Mints. *A Short Introduction to Modal Logic*. CSLI Publications, 1992.

[34] A. Plantinga. *The Nature of Necessity*. Oxford University Press, 1974.

[35] S. Popkorn. *First Steps in Modal Logic*. Cambridge University Press, 1992.

[36] G. Restall. Displaying and deciding substructural logics 1: Logics with contraposition. *Journal of Philosophical Logic*, 27:179–216, 1998.

[37] V. Shehtman and D. Skvortsov. Semantics of non-classical first-order predicate logics. In *Mathematical Logic*, pages 105–116. Plenum Press, 1990.

[38] S.K. Thomason. An incompleteness theorem in modal logic. *Theoria*, 40:150–158, 1974.

[39] H. Wansing. *Displaying Modal Logic*. Kluwer Academic Publishers, 1998.

[40] M. Zakharyashev, K. Segerberg, M. de Rijke, and H. Wansing. The origins of modern modal logic. In M. Zakharyashev et al., editor, *Advances in Modal Logic, Vol. 2*. CSLI Publications, 2000.

Maarten de Rijke
Institute for Logic, Language and Computation
University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
The Netherlands
E-mail: `mdr@science.uva.nl`
URL: `http://www.science.uva.nl/~mdr`

Heinrich Wansing
Institute of Philosophy
Dresden University of Technology
01062 Dresden
Germany
E-mail: `wansing@rcs1.tu-dresden.de`